Predictive inference for best linear combination of biomarkers subject to limits of detection

Tahani Coolen-Maturi*

Durham University Business School, Durham University, DH1 3LB

Abstract

Measuring the accuracy of diagnostic tests is crucial in many application areas including medicine, machine learning and credit scoring. The receiver operating characteristic (ROC) curve is a useful tool to assess the ability of a diagnostic test to discriminate between two classes or groups. In practice multiple diagnostic tests or biomarkers are combined to improve diagnostic accuracy. Often biomarker measurements are undetectable either below or above so-called limits of detection (LoD). In this paper, nonparametric predictive inference (NPI) for best linear combination of two or more biomarkers subject to limits of detection is presented. NPI is a frequentist statistical method that is explicitly aimed at using few modelling assumptions, enabled through the use of lower and upper probabilities to quantify uncertainty. The NPI lower and upper bounds for the ROC curve subject to limits of detection are derived, where the objective function to maximize is the area under the ROC curve (AUC). In addition, the paper discusses the effect of restriction on the linear combination's coefficients on the analysis. Examples are provided to illustrate the proposed method.

Keywords: Diagnostic accuracy; Limits of detection; Lower and upper probability; Imprecise probability; Nonparametric predictive inference; ROC curve.

1 Introduction

Measuring the accuracy of diagnostic tests is crucial in many application areas including medicine, machine learning and credit scoring. The receiver operating characteristic (ROC) curve is a useful tool to assess the ability of a diagnostic test to discriminate between two classes or groups. However, one diagnostic test may not be enough to draw a useful decision, thus in practice multiple diagnostic tests or biomarkers may be

*Email: tahani.maturi@durham.ac.uk (Corresponding author)

combined to improve diagnostic accuracy [1]. Another issue may occur when some biomarker measurements are undetectable either below or above some limits, so called limits of detection (LoD). Several papers considered combining biomarkers in order to improve the diagnostic accuracy, see e.g. [1, 2, 3, 4, 5]. Under the normality assumption, Su and Liu [6] considered a linear combination of biomarkers to maximize the area under the ROC curve while Liu *et al.* [7] extended their work to maximize the partial area under the ROC curve. Pepe and Thompson [1] introduced a logistic regression based approach to combine biomarkers. Perkins *et al.* [8, 9] generalized ROC curve inference for a single biomarker subject to a limit of detection. They also introduced a best linear combination of two biomarkers subject to limits of detection [10]. Dong *et al.* [11] proposed a method to estimate the multivariate normal distribution parameters taking the LoD into account, and then utilized the linear discriminant analysis to combine biomarkers. They have also described how to select and combine a subset of biomarkers based on the correlation distance to gain most accuracy. As with all parametric estimation methods, departure from the specified underlying distribution (e.g. the normality assumption) may lead to inaccurate estimation of the parameters [11]. On the other hand, nonparametric methods such as the one proposed in this paper provide an alternative regardless of the underlying distribution. For an extensive overview of the existing classical methods for combination of biomarkers, we refer the reader to Kang *et al.* [5] and Dong *et al.* [11]. These contributions to the literature are often either assuming some underlying distributions (e.g. normal) or focus on estimation rather than prediction. Prediction may be more natural in this context as one is typically interested in the performance (the accuracy) of diagnostic tests on detecting a specific condition for future patients. In this paper, nonparametric predictive inference (NPI) for best linear combination of two or more biomarkers subject to limits of detection is presented. NPI is a frequentist statistical method that is explicitly aimed at using few modelling assumptions, enabled through the use of lower and upper probabilities to quantify uncertainty [12, 13]. First, we briefly define the ROC curve and the area under this curve, AUC.

Let D be a binary variable describing the disease status, i.e. $D = 1$ for disease (cases) and $D = 0$ for non-disease (control). Suppose that X is a continuous random quantity of a diagnostic test result and that larger values of X are considered more indicative of disease. X^1 and X^0 are used to refer to test results for the disease and non-disease groups, respectively. The Receiver Operating Characteristic (ROC) curve is defined as the combination of False Positive Fraction (FPF) and True Positive Fraction (TPF) over all values of the threshold c, i.e. $\text{ROC} = \{(\text{FPF}(c), \text{TPF}(c)), c \in (-\infty, \infty)\}$, where $\text{FPF}(c) = P(X^0 > c)$ and $\text{TPF}(c) = P(X^1 > c)$. An ideal test completely separates the patients with and without the disease for a threshold c, i.e. $\text{FPF}(c) = 0$ and $\text{TPF}(c) = 1$. As the other extreme situation, if $\text{FPF}(c) = \text{TPF}(c)$ for all thresholds c, then the test has no ability to distinguish between the patients with and without the disease.

In many cases, a single numerical value or summary may be useful to represent the accuracy of a diagnostic

test or to compare two or more ROC curves [14]. A useful summary is the area under the ROC curve, $\text{AUC} = \int_0^1 \text{ROC}(t)\ dt$. The AUC measures the overall performance of the diagnostic test. Higher AUC values indicate more accurate tests, with AUC = 1 for perfect or ideal tests and AUC = 0.5 for uninformative tests. The AUC is equal to the probability that the test results from a randomly selected pair of diseased and non-diseased subjects are correctly ordered, i.e. $\text{AUC} = P\left[X^1 > X^0\right]$ [15]. So the AUC measures the test's ability to correctly classify a randomly selected individual as being from either the disease group or the non-disease group.

To estimate the ROC curve for diagnostic tests with continuous results, the nonparametric empirical method is popular due to its flexibility to adapt fully to the available data. This method yields the empirical ROC curve which will be considered, in particular to compare with the NPI method introduced in this paper. Suppose that we have test data on n_1 individuals from a disease group and n_0 individuals from a non-disease group, denoted by $\{x_i^1, i = 1, \ldots, n_1\}$ and $\{x_j^0, j = 1, \ldots, n_0\}$, respectively. Throughout this paper we assume that the two groups are fully independent, meaning that no information about any aspect related to one group contains information about any aspect of the other group. For the empirical method, these observations per group are assumed to be realisations of random quantities that are identically distributed as X^1 and X^0, for the disease and non-disease groups, respectively. The empirical estimator of the ROC is $\widehat{\text{ROC}} = \left\{\left(\widehat{\text{FPF}}(c), \widehat{\text{TPF}}(c)\right), c \in (-\infty, \infty)\right\}$ with $\widehat{\text{TPF}}(c) = \frac{1}{n_1}\sum_{i=1}^{n_1} \mathbf{1}\left\{x_i^1 > c\right\}$ and $\widehat{\text{FPF}}(c) = \frac{1}{n_0}\sum_{j=1}^{n_0} \mathbf{1}\left\{x_j^0 > c\right\}$, where $\mathbf{1}\{E\}$ is an indicator function which is equal to 1 if E is true and 0 else. The empirical estimator of the AUC is the well-known Mann-Whitney U statistic which is given by $\widehat{\text{AUC}} = \frac{1}{n_1 n_0}\sum_{j=1}^{n_0}\sum_{i=1}^{n_1}\left[\mathbf{1}\left\{x_i^1 > x_j^0\right\} + \frac{1}{2}\mathbf{1}\left\{x_i^1 = x_j^0\right\}\right]$ [14].

This paper is organized as follows. Section 2 provides a brief introduction to NPI. Section 3 presents NPI for one biomarker subject to limits of detection, followed by generalizing the results to find the best linear combination of two biomarkers subject to limits of detection in Section 4. Section 5 gives a matrix representation of the results in Sections 3 and 4 which simplifies the presentation of the proposed method when considering the general case of combining any finite number of biomarkers. Section 6 discusses the effect of restriction on the linear combination's coefficients in the analysis. Examples are provided to illustrate the proposed method in Section 7. The paper ends with some concluding remarks in Section 8 and an appendix presenting the proofs of the main results.

2 Nonparametric predictive inference

Nonparametric predictive inference (NPI) [12, 13] is based on the assumption $A_{(n)}$ proposed by Hill [16]. Let $X_1, \ldots, X_n, X_{n+1}$ be real-valued absolutely continuous and exchangeable random quantities. Let the ordered

observed values of $X_1, X_2, \ldots, X_n$ be denoted by $x_1 < x_2 < \ldots < x_n$ and let $x_0 = -\infty$ and $x_{n+1} = \infty$ for ease of notation. We assume that no ties occur; ties can be dealt with in NPI [13] but it is not relevant in this paper. For X_{n+1}, representing a future observation, $A_{(n)}$ [16] partially specifies a probability distribution by $P(X_{n+1} \in (x_{j-1}, x_j)) = \frac{1}{n+1}$ for $j = 1, \ldots, n+1$. $A_{(n)}$ does not assume anything else, and can be considered to be a post-data assumption related to exchangeability [17]. Inferences based on $A_{(n)}$ are predictive and nonparametric, and can be considered suitable if there is hardly any knowledge about the random quantity of interest, other than the n observations, or if one does not want to use such information. $A_{(n)}$ is not sufficient to derive precise probabilities for many events of interest, but it provides bounds for probabilities via the 'fundamental theorem of probability' [17], which are lower and upper probabilities in interval probability theory [18, 19, 20]. Augustin and Coolen [12] proved that NPI has strong consistency properties in the theory of interval probability. In NPI, uncertainty about the future observation X_{n+1} is quantified by lower and upper probabilities for events of interest. Lower and upper probabilities generalize classical ('precise') probabilities, and a lower (upper) probability for event A, denoted by $\underline{P}(A)$ ($\overline{P}(A)$), can be interpreted as supremum buying (infimum selling) price for a gamble on the event A [18], or just as the maximum lower (minimum upper) bound for the probability of A that follows from the assumptions made [13]. This latter interpretation is used in NPI, we wish to explore application of $A_{(n)}$ for inference without making further assumptions. So, NPI lower and upper probabilities are the sharpest bounds on a probability for an event of interest when only $A_{(n)}$ is assumed. Informally, $\underline{P}(A)$ ($\overline{P}(A)$) can be considered to reflect the evidence in favour of (against) event A.

NPI has been introduced for assessing the accuracy of a classifier's ability to discriminate between two outcomes (or two groups) for binary data [21] and for diagnostic tests with ordinal observations [22] and with real-valued observations [23]. Recently, [24] generalized the results in [23] by introducing NPI for three-group ROC analysis, with real-valued observations, to assess the ability of a diagnostic test to discriminate among three ordered classes or groups. Coolen-Maturi [25] generalized the results in [22] by proposing NPI for three-group ROC analysis with ordinal outcomes. Below we give a brief overview of NPI for two-group ROC analysis following [23].

Suppose that $\{X_i^1,\ i = 1, \ldots, n_1, n_1 + 1\}$ are continuous and exchangeable random quantities from the disease group and $\{X_j^0,\ j = 1, \ldots, n_0, n_0 + 1\}$ are continuous and exchangeable random quantities from the non-disease group, where $X_{n_1+1}^1$ and $X_{n_0+1}^0$ are the next observations from the disease and non-disease groups following n_1 and n_0 observations, respectively. As mentioned before, we assume that both groups are fully independent. Let $x_1^1 < \ldots < x_{n_1}^1$ be the ordered observed values for the first n_1 individuals from the disease group and $x_1^0 < \ldots < x_{n_0}^0$ the ordered observed values for the first n_0 individuals from the non-disease group. For ease of notation, let $x_0^1 = x_0^0 = -\infty$ and $x_{n_1+1}^1 = x_{n_0+1}^0 = \infty$. We assume that there

are no ties in the data, it can be easily generalized to allow ties [13]. The NPI lower and upper ROC curves are $\underline{\text{ROC}} = \left\{\left(\overline{\text{FPF}}(c), \underline{\text{TPF}}(c)\right), c \in (-\infty, \infty)\right\}$ and $\overline{\text{ROC}} = \left\{\left(\underline{\text{FPF}}(c), \overline{\text{TPF}}(c)\right), c \in (-\infty, \infty)\right\}$ [23], where

$$\underline{TPF}(c) = \underline{P}(X^1_{n_1+1} > c) = \frac{\sum_{i=1}^{n_1} \mathbf{1}\left\{x_i^1 > c\right\}}{n_1+1}, \quad \overline{TPF}(c) = \overline{P}(X^1_{n_1+1} > c) = \frac{\sum_{i=1}^{n_1} \mathbf{1}\left\{x_i^1 > c\right\}+1}{n_1+1},$$
$$\underline{FPF}(c) = \underline{P}(X^0_{n_0+1} > c) = \frac{\sum_{j=1}^{n_0} \mathbf{1}\left\{x_j^0 > c\right\}}{n_0+1}, \quad \overline{FPF}(c) = \overline{P}(X^0_{n_0+1} > c) = \frac{\sum_{j=1}^{n_0} \mathbf{1}\left\{x_j^0 > c\right\}+1}{n_0+1},$$

and $\underline{P}$ and $\overline{P}$ are NPI lower and upper probabilities [12]. It is easily seen that $\underline{\text{FPF}}(c) \leq \widehat{\text{FPF}}(c) \leq \overline{\text{FPF}}(c)$ and $\underline{\text{TPF}}(c) \leq \widehat{\text{TPF}}(c) \leq \overline{\text{TPF}}(c)$ for all c, which implies that the empirical ROC curve is bounded by the NPI lower and upper ROC curves [23].

The NPI lower and upper AUC are defined as the NPI lower and upper probabilities for the event that the test result for the next individual from the disease group is greater than the test result for the next individual from the non-disease group, as given by [23]

$$\underline{\text{AUC}} = \underline{P}\left(X^1_{n_1+1} > X^0_{n_0+1}\right) = \frac{1}{(n_1+1)(n_0+1)} \sum_{j=1}^{n_0} \sum_{i=1}^{n_1} \mathbf{1}\left\{x_i^1 > x_j^0\right\}, \tag{1}$$

$$\overline{\text{AUC}} = \overline{P}\left(X^1_{n_1+1} > X^0_{n_0+1}\right) = \frac{1}{(n_1+1)(n_0+1)} \left[\sum_{j=1}^{n_0} \sum_{i=1}^{n_1} \mathbf{1}\left\{x_i^1 > x_j^0\right\} + n_1 + n_0 + 1\right]. \tag{2}$$

It is interesting to notice that the imprecision $\overline{\text{AUC}} - \underline{\text{AUC}} = \frac{n_1+n_0+1}{(n_1+1)(n_0+1)}$ depends only on the two sample sizes n_0 and n_1. Coolen-Maturi *et al.* [23] showed that equation (1) is actually the area under the $\underline{\text{ROC}}$ and equation (2) is the area under the $\overline{\text{ROC}}$.

3 Predictive inference for a single biomarker subject to limits of detection

Biomarker measurements may be subject to limits of detection, e.g. due to instrumental limitation measurements may be undetectable below or above certain limits. Perkins *et al.* [8] showed that ignoring these measurements or even plugging in some replacement values can lead to biased estimates of the area under the curve. In this section, NPI for ROC curve and the area under the ROC curve, AUC, for a single biomarker subject to limits of detection are presented. The proposed method provides an alternative approach to treat this issue, namely in our proposed NPI method the measurements that are below or above the limits of detections are not removed or replaced by other values, instead only their numbers are taken into account to derive the NPI lower and upper ROC curves and the area under these curves. We will show later via exam-

ples in Section 7 how the proposed method captures the issue of the limits of detection and how it quantifies the uncertainty via the lower and upper probabilities. But first we need to introduce some notation.

Let X be a biomarker whose measurements are affected by limits of detection $L_x < U_x$, for each group these limits $L_x < U_x$ divide the data into three parts. For the disease (non-disease) group, there are l_x^1 (l_x^0) observations which are only known to be less than L_x, u_x^1 (u_x^0) which are only known to be greater than U_x, while the r_x^1 (r_x^0) ordered observations between L_x and U_x are fully known and denoted by

$$
\begin{aligned}
&-\infty < L_x \leq x_{(1)}^1 < x_{(2)}^1 < \ldots < x_{(r_x^1)}^1 \leq U_x < \infty, \\
&-\infty < L_x \leq x_{(1)}^0 < x_{(2)}^0 < \ldots < x_{(r_x^0)}^0 \leq U_x < \infty.
\end{aligned}
$$

For ease of presentation, let $x_{(0)}^1 = x_{(0)}^0 = -\infty$ and $x_{(r_x^1+1)}^1 = x_{(r_x^0+1)}^0 = \infty$. We should mention here that $-\infty$ and ∞ are just indicators of the range of possible values for $X_{n_0+1}^0$ and $X_{n_1+1}^1$, e.g. if biomarker results are believed to be only positive numbers then one would set $x_{(0)}^1 = x_{(0)}^0 = 0$.

As we have limits of detection, we cannot use the $A_{(n)}$ assumption directly to derive the NPI lower and upper ROC curves. Therefore, we need to use the generalized $A_{(n)}^{tt}$ assumption introduced by Maturi *et al.* [26] for terminated data. The following theorem describes how the probability distribution for X_{n+1} is partially specified by so-called M-functions. We should mention here that the intervals on which M-functions are defined can be overlapped and all M-function values must sum up to one. The concept of M-function is similar to that of Shafer's basic probability assignment [27].

Theorem 1. *The assumption $A_{(n)}^{tt}$ is that the probability distribution for a real-valued random quantity X_{n+1}, on the basis of the data terminated at two cut points L_x and U_x as described above, is partially specified by the following M-function values:*

$$
\begin{aligned}
&M_{X_{n+1}}(x_{(i)}, x_{(i+1)}) = \frac{1}{n+1} \quad , \quad i = 0, 1, \ldots, r_x, \\
&M_{X_{n+1}}(-\infty, L_x) = \frac{l_x}{n+1} \qquad \textit{and} \qquad M_{X_{n+1}}(U_x, \infty) = \frac{u_x}{n+1}.
\end{aligned}
$$

By applying the assumption $A_{(n)}^{tt}$ per group, i.e. $A_{(n_1)}^{tt}$ for the disease group and $A_{(n_0)}^{tt}$ for the non-disease

group, we can derive the lower and upper bounds for FPF and TPF as follows:

$$\underline{FPF}(c) = \underline{P}(X^0_{n_0+1} > c) = \frac{1}{n_0+1}\left[\sum_{i=1}^{r^0_x} \mathbf{1}\{x^0_{(i)} > c\} + \mathbf{1}\{U_x > c\}\, u^0_x\right],$$

$$\overline{FPF}(c) = \overline{P}(X^0_{n_0+1} > c) = \frac{1}{n_0+1}\left[\sum_{i=1}^{r^0_x} \mathbf{1}\{x^0_{(i)} > c\} + \mathbf{1}\{L_x > c\}\, l^0_x + u^0_x + 1\right],$$

$$\underline{TPF}(c) = \underline{P}(X^1_{n_1+1} > c) = \frac{1}{n_1+1}\left[\sum_{i=1}^{r^1_x} \mathbf{1}\{x^1_{(i)} > c\} + \mathbf{1}\{U_x > c\}\, u^1_x\right],$$

$$\overline{TPF}(c) = \overline{P}(X^1_{n_1+1} > c) = \frac{1}{n_1+1}\left[\sum_{i=1}^{r^1_x} \mathbf{1}\{x^1_{(i)} > c\} + \mathbf{1}\{L_x > c\}\, l^1_x + u^1_x + 1\right].$$

The NPI lower and upper ROC curves can be defined as

$$\underline{ROC} = \{(\overline{FPF}(c), \underline{TPF}(c)), c \in (-\infty, \infty)\}, \tag{3}$$

$$\overline{ROC} = \{(\underline{FPF}(c), \overline{TPF}(c)), c \in (-\infty, \infty)\}. \tag{4}$$

If all the biomarker measurements are observed, i.e. $r^0_x = n^0$ and $r^1_x = n^1$ [hence $l^0_x = u^0_x = l^1_x = u^1_x = 0$] then we get the complete data case presented by Coolen-Maturi *et al.* [23].

The areas under the lower and upper ROC curves (AUC) are defined as the lower and upper probabilities for the event $X^0_{n_0+1} < X^1_{n_1+1}$ [23]. As the biomarker test results are subject to limits of detection, we cannot use the results in [23]. Maturi *et al.* [26] introduced NPI for comparing two groups of real-valued data with terminated tails, where we only know the number of observations beyond the terminated points. In this paper, we are going to utilize the results in [26] for the areas under the lower and upper ROC curves by using the M-functions introduced in Theorem 1. The areas under the lower and upper ROC curves in (3) and (4) are given by the following theorem.

Theorem 2. *The areas under the lower and upper ROC curves for a biomarker X, subject to limits of detection, are*

$$\underline{AUC} = \underline{P}(X^0_{n_0+1} < X^1_{n_1+1}) = \frac{1}{(n_0+1)(n_1+1)}\left[\sum_{j=1}^{r^1_x}\sum_{i=1}^{r^0_x} \mathbf{1}\{x^0_{(i)} < x^1_{(j)}\} + l^0_x(r^1_x + u^1_x) + r^0_x u^1_x\right], \tag{5}$$

$$\overline{AUC} = \overline{P}(X^0_{n_0+1} < X^1_{n_1+1}) = \frac{1}{(n_0+1)(n_1+1)}\left[\sum_{j=1}^{r^1_x}\sum_{i=1}^{r^0_x} \mathbf{1}\{x^0_{(i)} < x^1_{(j)}\} + (l^0_x+1)(l^1_x + r^1_x) + (u^1_x+1)(n_0+1)\right]. \tag{6}$$

The imprecision is the difference between $\overline{AUC}$ and $\underline{AUC}$ which reflects the amount of information available.

4 Predictive inference for best linear combination of two biomarkers subject to limits of detection

In medical applications, researchers may want to combine two biomarkers to improve diagnostic accuracy. Improving the diagnostic accuracy is often done by maximizing the area under the ROC curve. In this setting each subject has two biomarker measurements, where the two biomarkers, say X and Y, may be subject to limits of detection. In this section, we extend the approach presented in Section 3 to combine two biomarkers, subject to limits of detection, to improve the accuracy by maximizing the lower and upper AUC. In addition to the notation introduced in Section 3 we need to introduce further notation as follows.

Suppose that $\{Y_i^1, i = 1, \ldots, n_1, n_1 + 1\}$ are continuous and exchangeable random quantities from the disease group and $\{Y_j^0, j = 1, \ldots, n_0, n_0 + 1\}$ are continuous and exchangeable random quantities from the non-disease group, where $Y_{n_1+1}^1$ and $Y_{n_0+1}^0$ are the next observations from the disease and non-disease groups, respectively. Let Y be a biomarker whose measurements are affected by limits of detection $L_y < U_y$, for each group these limits $L_y < U_y$ divide the data into three parts. For the disease (non-disease) group, there are l_y^1 (l_y^0) observations which are only known to be less than L_y, u_y^1 (u_y^0) which are only known to be greater than U_y, while the r_y^1 (r_y^0) ordered observations between L_y and U_y are fully known and denoted by

$$
\begin{aligned}
-\infty < L_y \leq y_{(1)}^1 < y_{(2)}^1 < \ldots < y_{(r_y^1)}^1 \leq U_y < \infty, \\
-\infty < L_y \leq y_{(1)}^0 < y_{(2)}^0 < \ldots < y_{(r_y^0)}^0 \leq U_y < \infty.
\end{aligned}
$$

For ease of presentation, let $y_{(0)}^1 = y_{(0)}^0 = -\infty$ and $y_{(r_y^1+1)}^1 = y_{(r_y^0+1)}^0 = \infty$. We can define the lower and upper ROC curves and the corresponding areas under these curves for biomarker Y as in Section 3. Recall that, for biomarker X (Y), $X_{n_1+1}^1$ and $X_{n_0+1}^0$ ($Y_{n_1+1}^1$ and $Y_{n_0+1}^0$) are the next future observations from the disease and non-disease group, respectively. Now we are interested in combining these future observations by defining the scores $T_{n_0+1}^0 = \alpha_1 X_{n_0+1}^0 + \alpha_2 Y_{n_0+1}^0$ and $T_{n_1+1}^1 = \alpha_1 X_{n_1+1}^1 + \alpha_2 Y_{n_1+1}^1$. The objective is to find the values of α_1 and α_2 that maximize the lower and upper areas under the ROC curves (i.e. maximize the accuracy).

In order to introduce the corresponding M-functions for $T_{n+1} = \alpha_1 X_{n+1} + \alpha_2 Y_{n+1}$, we have dropped the superscripts here for ease of presentation, we need to define further notation. Given the four limits of detection points L_x, L_y, U_x and U_y, the data structure with these four limits can be visualized as in Figure 1,

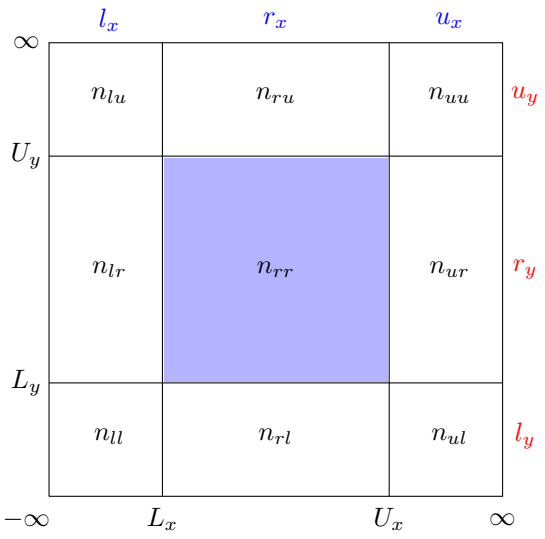

Figure 1: Data structure

where e.g. n_{ll} is the number of individuals whose X and Y measurements are below L_x and L_y, respectively. Other quantities are defined similarly,

$$
\begin{aligned}
&n_{ll} = \#\{x < L_x \wedge y < L_y\}, && n_{rl} = \#\{L_x < x < U_x \wedge y < L_y\}, && n_{ul} = \#\{x > U_x \wedge y < L_y\},\\
&n_{lr} = \#\{x < L_x \wedge L_y < y < U_y\}, && n_{rr} = \#\{L_x < x < U_x \wedge L_y < y < U_y\}, && n_{ur} = \#\{x > U_x \wedge L_y < y < U_y\},\\
&n_{lu} = \#\{x < L_x \wedge y > U_y\}, && n_{ru} = \#\{L_x < x < U_x \wedge y > U_y\}, && n_{uu} = \#\{x > U_x \wedge y > U_y\}.
\end{aligned}
$$

This leads to the equalities $l_x = n_{ll} + n_{lr} + n_{lu}$, $r_x = n_{rl} + n_{rr} + n_{ru}$ and $u_x = n_{ul} + n_{ur} + n_{uu}$, and similarly for l_y, r_y and u_r, and $n = l_x + r_x + u_x = l_y + r_y + u_y$. The data structure in Figure 1 can also be expressed in a matrix format as

$$
\mathbb{S} = \begin{array}{c} \begin{matrix} l_x & r_x & u_x \end{matrix} \\ \begin{pmatrix} n_{lu} & n_{ru} & n_{uu} \\ n_{lr} & n_{rr} & n_{ur} \\ n_{ll} & n_{rl} & n_{ul} \end{pmatrix} \end{array} \begin{matrix} u_y \\ r_y \\ l_y \end{matrix}.
$$

Let $r_T = n_{rr}$ be the number of the observed values (available data) from both biomarkers X and Y, that is the test results of biomarker X that are between L_x and U_x and the test results of biomarker of Y that are between L_y and U_y. Thus, the combined test results from both biomarkers are $t_i = \alpha_1 x_i + \alpha_2 y_i$, $i = 1, \ldots, r_T$ where $r_T = n_{rr}$. Let $t_{(i)}$ be the ith value among t_i, thus $-\infty < \alpha_1 L_x + \alpha_2 L_y < t_{(1)} < \ldots < t_{(r_T)} < \alpha_1 U_x + \alpha_2 U_y < \infty$. The probability mass specifications for $T_{n+1} = \alpha_1 X_{n+1} + \alpha_2 Y_{n+1}$ are given by the M-functions in Definition 1. This generalizes the assumption $A^{tt}_{(n)}$ in Theorem 1, we will denote the new

generalized assumption as $A^{tt2}_{(n)}$.

Definition 1 (Assumption $A^{tt2}_{(n)}$)**.** *The assumption $A^{tt2}_{(n)}$ is that the probability distribution for a real-valued random quantity $T_{n+1} = \alpha_1 X_{n+1} + \alpha_2 Y_{n+1}$, on the basis of the data terminated at the two cut points L_x and U_x for X, and at the two cut points L_y and U_y for Y as described above, is partially specified by the following M-function values:*

$$M_{T_{n+1}}(t_{(i)}, t_{(i+1)}) = \frac{1}{n+1} \ , \ \ i = 0, 1, \ldots, r_T,$$

where $r_T = n_{rr}$, $t_{(0)} = -\infty$ and $t_{(r_T+1)} = \infty$. Notice that $\sum_{i=0}^{r_T} M_{T_{n+1}}(t_{(i)}, t_{(i+1)}) = \frac{r_T+1}{n+1} = \frac{n_{rr}+1}{n+1}$. The remaining M-functions are defined as

$$\begin{aligned}
M_{T_{n+1}}(X_{n+1} \in (-\infty, L_x), Y_{n+1} \in (-\infty, L_y)) &= M_{T_{n+1}}(-\infty, \alpha_1 L_x + \alpha_2 L_y) = \frac{n_{ll}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (-\infty, L_x), Y_{n+1} \in (L_y, U_y)) &= M_{T_{n+1}}(-\infty, \alpha_1 L_x + \alpha_2 U_y) = \frac{n_{lr}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (-\infty, L_x), Y_{n+1} \in (U_y, \infty)) &= M_{T_{n+1}}(-\infty, \infty) = \frac{n_{lu}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (L_x, U_x), Y_{n+1} \in (-\infty, L_y)) &= M_{T_{n+1}}(-\infty, \alpha_1 U_x + \alpha_2 L_y) = \frac{n_{rl}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (L_x, U_x), Y_{n+1} \in (U_y, \infty)) &= M_{T_{n+1}}(\alpha_1 L_x + \alpha_2 U_y, \infty) = \frac{n_{ru}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (U_x, \infty), Y_{n+1} \in (-\infty, L_y)) &= M_{T_{n+1}}(-\infty, \infty) = \frac{n_{ul}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (U_x, \infty), Y_{n+1} \in (L_y, U_y)) &= M_{T_{n+1}}(\alpha_1 U_x + \alpha_2 L_y, \infty) = \frac{n_{ur}}{n+1}, \\
M_{T_{n+1}}(X_{n+1} \in (U_x, \infty), Y_{n+1} \in (U_y, \infty)) &= M_{T_{n+1}}(\alpha_1 U_x + \alpha_2 U_y, \infty) = \frac{n_{uu}}{n+1}.
\end{aligned}$$

By applying the assumption $A^{tt2}_{(n)}$ per group, i.e. $A^{tt2}_{(n_0)}$ and $A^{tt2}_{(n_1)}$, we can define the NPI lower and upper ROC curves as

$$\underline{ROC} = \{(\overline{FPF}(c), \underline{TPF}(c)), c \in (-\infty, \infty)\}, \tag{7}$$

$$\overline{ROC} = \{(\underline{FPF}(c), \overline{TPF}(c)), c \in (-\infty, \infty)\}. \tag{8}$$

where

$$\begin{aligned}
\underline{FPF}(c) =& \frac{1}{n_0+1} \sum_{j=1}^{r_T^0} \mathbf{1}\{t^0_{(j)} > c\} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y > c\} \frac{n^0_{ru}}{n_0+1} \\
&+ \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y > c\} \frac{n^0_{ur}}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y > c\} \frac{n^0_{uu}}{n_0+1}.
\end{aligned} \tag{9}$$

$$\overline{FPF}(c) = \frac{1}{n_0+1}\sum_{j=1}^{r_T^0} \mathbf{1}\{t_{(j)}^0 > c\} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y > c\}\frac{n_{ll}^0}{n_0+1} + \frac{1+u_y^0+u_x^0-n_{uu}^0}{n_0+1}$$
$$+\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y > c\}\frac{n_{lr}^0}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y > c\}\frac{n_{rl}^0}{n_0+1}. \tag{10}$$

$$\underline{TPF}(c) = \frac{1}{n_1+1}\sum_{i=1}^{r_T^1} \mathbf{1}\{t_{(i)}^1 > c\} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y > c\}\frac{n_{ru}^1}{n_1+1}$$
$$+\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y > c\}\frac{n_{ur}^1}{n_1+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y > c\}\frac{n_{uu}^1}{n_1+1}. \tag{11}$$

$$\overline{TPF}(c) = \frac{1}{n_1+1}\sum_{i=1}^{r_T^1} \mathbf{1}\{t_{(i)}^1 > c\} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y > c\}\frac{n_{ll}^1}{n_1+1} + \frac{1+u_y^1+u_x^1-n_{uu}^1}{n_1+1}$$
$$+\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y > c\}\frac{n_{lr}^1}{n_1+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y > c\}\frac{n_{rl}^1}{n_1+1}. \tag{12}$$

The lower and upper bounds for the area under these lower and upper ROC curves, $\underline{AUC}$ and $\overline{AUC}$, are given by Theorem 3, which is equivalent to finding the lower and upper probabilities for the event $T_{n_0+1}^0 < T_{n_1+1}^1$.

Theorem 3. *The NPI lower and upper bounds for the area under the ROC curves, $\underline{AUC}$ and $\overline{AUC}$, which are equal to the lower and upper probabilities for the event $T_{n_0+1}^0 < T_{n_1+1}^1$, are given by*

$$\underline{AUC} = \underline{P}(T_{n_0+1}^0 < T_{n_1+1}^1) = \sum_{i=1}^{r_T^1}\frac{1}{n_1+1}\left[\frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t_{(i)}^1\}\frac{n_{lr}^0}{n_0+1}\right.$$
$$\left.+\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t_{(i)}^1\}\frac{n_{rl}^0}{n_0+1} + \sum_{j=0}^{r_T^0-1}\mathbf{1}\{t_{(j+1)}^0 < t_{(i)}^1\}\frac{1}{n_0+1}\right] + \frac{n_{uu}^1}{n_1+1}\left[\frac{n_{ll}^0+n_{lr}^0+n_{rl}^0+n_{rr}^0}{n_0+1}\right]$$
$$+\frac{n_{ru}^1}{n_1+1}\left[\frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n_{rl}^0}{n_0+1} + \sum_{j=0}^{r_T^0-1}\mathbf{1}\{t_{(j+1)}^0 < \alpha_1 L_x + \alpha_2 U_y\}\frac{1}{n_0+1}\right]$$
$$+\frac{n_{ur}^1}{n_1+1}\left[\frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n_{lr}^0}{n_0+1} + \sum_{j=0}^{r_T^0-1}\mathbf{1}\{t_{(j+1)}^0 < \alpha_1 U_x + \alpha_2 L_y\}\frac{1}{n_0+1}\right],$$

$$\overline{AUC} = \overline{P}(T_{n_0+1}^0 < T_{n_1+1}^1) = \frac{u_x^1+u_y^1-n_{uu}^1+1}{n_1+1} + \left[\frac{n_{rr}^1+n_{ll}^1+n_{lr}^1+n_{rl}^1}{n_1+1}\right]\left\{\frac{l_x^0+l_y^0-n_{ll}^0+1}{n_0+1}\right\}$$
$$+\sum_{i=1}^{r_T^1}\frac{1}{n_1+1}\left[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t_{(i)}^1\}\frac{n_{ru}^0}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t_{(i)}^1\}\frac{n_{ur}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < t_{(i)}^1\}\frac{1}{n_0+1}\right]$$
$$+\frac{n_{lr}^1}{n_1+1}\left[\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n_{ur}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < \alpha_1 L_x + \alpha_2 U_y\}\frac{1}{n_0+1}\right]$$
$$+\frac{n_{rl}^1}{n_1+1}\left[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n_{ru}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < \alpha_1 U_x + \alpha_2 L_y\}\frac{1}{n_0+1}\right].$$

The proof of Theorem 3 is given in Appendix A1.

When all the biomarker measurements are observed, i.e. $r_T^1 = n_{rr}^1 = n_1$ and $r_T^0 = n_{rr}^0 = n_0$ [hence $l_x^0 = u_x^0 = l_x^1 = u_x^1 = 0$ and $l_y^0 = u_y^0 = l_y^1 = u_y^1 = 0$], then equations (9)-(12) are equal to

$$\underline{FPF}(c) = \underline{P}(T_{n_0+1}^0 > c) = \frac{\sum_{j=1}^{n_0} \mathbf{1}\{t_{(j)}^0 > c\}}{n_0+1} \quad , \quad \overline{FPF}(c) = \overline{P}(T_{n_0+1}^0 > c) = \frac{\sum_{j=1}^{n_0} \mathbf{1}\{t_{(j)}^0 > c\} + 1}{n_0+1}.$$
$$\underline{TPF}(c) = \underline{P}(T_{n_1+1}^1 > c) = \frac{\sum_{i=1}^{n_1} \mathbf{1}\{t_{(i)}^1 > c\}}{n_1+1} \quad , \quad \overline{TPF}(c) = \overline{P}(T_{n_1+1}^1 > c) = \frac{\sum_{i=1}^{n_1} \mathbf{1}\{t_{(i)}^1 > c\} + 1}{n_1+1}.$$

and the NPI lower and upper bounds for the area under the ROC curve, that are given in Theorem 3, can be written as

$$\underline{AUC} = \underline{P}(T_{n_0+1}^0 < T_{n_1+1}^1) = \frac{1}{(n_0+1)(n_1+1)} \sum_{i=1}^{n_1} \sum_{j=1}^{n_0} \mathbf{1}\{t_{(j)}^0 < t_{(i)}^1\},$$
$$\overline{AUC} = \overline{P}(T_{n_0+1}^0 < T_{n_1+1}^1) = \frac{1}{(n_0+1)(n_1+1)} \left[\sum_{i=1}^{n_1} \sum_{j=1}^{n_0} \mathbf{1}\{t_{(j)}^0 < t_{(i)}^1\} + n_1 + n_0 + 1 \right].$$

which is equivalent to applying the approach proposed by Coolen-Maturi *et al.* [23] on the combined scores.

The question now is how to find the values $\underline{\alpha} = (\alpha_1, \alpha_2)$ that maximize the NPI lower and upper AUC. Our proposed method does not impose any restriction on the values of α_1 and α_2. For example, one could maximize the AUC by maximizing $\underline{\alpha} = (1, \alpha)$, where $\alpha = \alpha_2/\alpha_1$, which is similar to the way used by Pepe and Thompson [1]. According to Pepe and Thompson [1], this maximization is implemented by searching α, in which the area under the curve AUC corresponding to the combined test $X + \alpha Y$ is evaluated for 201 equally spaced values of $\alpha \in [-1, 1]$. For $\alpha < -1$ and $\alpha > 1$, $AUC(\gamma X + Y)$ where $\gamma = \frac{1}{\alpha} \in [-1, 1]$, thus the AUC corresponding to the combined test $\gamma X + Y$, is evaluated for another 201 equally spaced values of $\gamma = \frac{1}{\alpha} \in [-1, 1]$. The optimal combination coefficient is $\underline{\alpha} = (1, \alpha)$ or $\underline{\alpha} = (\gamma, 1)$ that maximizes the AUC. As mentioned above, the proposed method in this paper can be used without any restriction on the values of α_1 and α_2, however, for the examples in this paper, we propose to find the $\underline{\alpha} = (\alpha_1, \alpha_2)$ that maximizes the NPI lower and upper AUC such that $\alpha_1, \alpha_2 \in [0, 1]$ and $\alpha_1 + \alpha_2 = 1$. The maximization is implemented by searching α_1 and α_2, the AUC corresponding to the combined test $\alpha_1 X + \alpha_2 Y$ (or $\alpha_1 X_{n+1} + \alpha_2 Y_{n+1}$ for NPI) is evaluated for 101 equally spaced values for each $\alpha_1 \in [0, 1]$ and $\alpha_2 \in [0, 1]$ such that $\alpha_1 + \alpha_2 = 1$. More discussion about the use of this restriction and its advantages is given in Section 6.

Before we illustrate our method via examples, we should look at the issue of data preparation and processing. As we would like to combine two biomarkers, we should make sure that biomarker measurements are comparable, e.g. have the same units and same value range, otherwise we need to rescale or normalize the

data. There are many normalization methods available, e.g. to scale the data to be between any arbitrary points a and b we can use

$$\tilde{x} = a\left(\frac{x_{\max} - x}{x_{\max} - x_{\min}}\right) + b\left(\frac{x - x_{\min}}{x_{\max} - x_{\min}}\right). \tag{13}$$

or we can standardize the data, using the mean $\overline{x}$ and the standard deviation s as $\tilde{x} = \frac{x-\overline{x}}{s}$. Alternatively, a normalization method that is more robust against outliers is derived at by using $\tilde{x} = \frac{x-Q_2}{Q_3-Q_1}$, where Q_1, Q_2 and Q_3 are the first, the second (median) and the third quartiles, leading to normalized values with median zero and interquartile range (IQR) equal to one.

The results presented in this section implicitly assumed that the two biomarker tests results are comparable or that some kind of normalization has been carried out. For example, one could normalize the biomarker tests results (the disease and non-disease results combined per biomarker) using equation (13) or replacing the min and the max by L_x and U_x instead, respectively. In the case of a single biomarker, i.e. using the analysis presented in Section 3, we get the same results (lower and upper AUC) whether we normalize the data or not (or e.g. taking the log or not). In fact we get the same results if we apply any monotone function.

5 Results in matrix formulation

So far we have presented NPI for combining two biomarkers, our method can be easily extended for combining more than two biomarkers as we will discuss below, but first we need to present the results introduced in previous sections in matrix formulation. This matrix representation is particularly useful in presenting and implementing (in statistical software such as R [28]) the method for combining several biomarkers.

The NPI lower and upper probabilities in Theorem 3 can be written in a matrix format as follows: Let Q_L and Q_U be two matrices of order $(r_T^{D=0}+9)\times(r_T^{D=1}+9)$, where $D=1$ and $D=0$ refer to the disease and non-disease group, respectively, such that

$$Q_L = q[j,i] = \begin{cases} 1 & \text{if} \quad IU_j^{D=0} < IL_i^{D=1} \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad Q_U = q[j,i] = \begin{cases} 1 & \text{if} \quad IL_j^{D=0} < IU_i^{D=1} \\ 0 & \text{otherwise} \end{cases},$$

where IL^D and IU^D are two vectors of order (r_T^D+9) consisting of the lower- and the upper-end limits of the intervals in Definition 1, respectively. Let M^D be a vector of order (r_T^D+9) consisting of the probability

mass functions (M-functions) corresponding to these intervals, that is

$$IL^D = \begin{bmatrix} t^D_{(0)} \\ t^D_{(1)} \\ \vdots \\ t^D_{(r^D_T-1)} \\ t^D_{(r^D_T)} \\ \cdots\cdots\cdots\cdots \\ -\infty \\ -\infty \\ -\infty \\ -\infty \\ \alpha_1 L_x + \alpha_2 U_y \\ -\infty \\ \alpha_1 U_x + \alpha_2 L_y \\ \alpha_1 U_x + \alpha_2 U_y \end{bmatrix}, \qquad IU^D = \begin{bmatrix} t^D_{(1)} \\ t^D_{(2)} \\ \vdots \\ t^D_{(r^D_T)} \\ t^D_{(r^D_T+1)} \\ \cdots\cdots\cdots\cdots \\ \alpha_1 L_x + \alpha_2 L_y \\ \alpha_1 L_x + \alpha_2 U_y \\ \infty \\ \alpha_1 U_x + \alpha_2 L_y \\ \infty \\ \infty \\ \infty \\ \infty \end{bmatrix}, \qquad M^D = \frac{1}{n_D+1} \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ \cdots\cdots \\ n^D_{ll} \\ n^D_{lr} \\ n^D_{lu} \\ n^D_{rl} \\ n^D_{ru} \\ n^D_{ul} \\ n^D_{ur} \\ n^D_{uu} \end{bmatrix}.$$

The NPI lower and upper bounds for the area under the ROC curves, in Theorem 3, can be written as

$$\underline{AUC} = \underline{P}(T^0_{n_0+1} < T^1_{n_1+1}) = \left(M^{D=0}\right)^{tr} Q_L M^{D=1},$$
$$\overline{AUC} = \overline{P}(T^0_{n_0+1} < T^1_{n_1+1}) = \left(M^{D=0}\right)^{tr} Q_U M^{D=1}.$$

where $(.)^{tr}$ refers to the transpose of a matrix. The lower and upper ROC curves can be written in matrix format as

$$\underline{FPF}(c) = \mathbf{I}_{\{IL^{D=0}_j > c\}} \left(M^{D=0}\right)^{tr} \quad , \quad \overline{FPF}(c) = \mathbf{I}_{\{IU^{D=0}_j > c\}} \left(M^{D=0}\right)^{tr},$$
$$\underline{TPF}(c) = \mathbf{I}_{\{IL^{D=1}_i > c\}} \left(M^{D=1}\right)^{tr} \quad \text{and} \quad \overline{TPF}(c) = \mathbf{I}_{\{IU^{D=1}_i > c\}} \left(M^{D=1}\right)^{tr}.$$

where $\mathbf{I}_{\{A_j\}}$ is an indicator vector of order (r^D_T+9) whose jth element is equal to 1 if A_j is true and 0 otherwise.

The NPI lower and upper ROC curves and the areas under these curves (AUC) for a single biomarker, as introduced in Section 3, can also be written in a matrix format as above, where

$$IL^D = \begin{bmatrix} x^D_{(0)} \\ x^D_{(1)} \\ \vdots \\ x^D_{(r^D_x-1)} \\ x^D_{(r^D_x)} \\ \cdots\cdots \\ -\infty \\ U_x \end{bmatrix}, \qquad IU^D = \begin{bmatrix} x^D_{(1)} \\ x^D_{(2)} \\ \vdots \\ x^D_{(r^D_x)} \\ x^D_{(r^D_x+1)} \\ \cdots\cdots \\ L_x \\ \infty \end{bmatrix}, \qquad M^D = \frac{1}{n_D+1} \begin{bmatrix} ,1 \\ 1 \\ \vdots \\ 1 \\ 1 \\ \cdots \\ l^D_x \\ u^D_x \end{bmatrix}.$$

In this case, Q_L and Q_U are matrices of order $(r^0_x+3)\times(r^1_x+3)$ and they are defined as above.

This matrix representation is particularly useful in presenting the method for combining more than two biomarkers. In general, if we have K biomarkers and we are interested in combining them to improve the accuracy, that is $T_{n+1}=\sum_{j=1}^K \alpha_j Z_{j,n+1}$, with $\alpha_j\in[0,1]$ and $\sum_{j=1}^K\alpha_j=1$, then the two matrices Q_L and Q_U will be of order $(r_T^{D=0}+3^K)(r_T^{D=1}+3^K)$ and we need to define the vector M^D of order $(r_T^D+3^K)$, the M-functions, as we did in Definition 1.

Let $r_T=n_{rr\ldots r}$ be the number of the observed values (available data) from all K biomarkers, that is the test results of biomarker Z_i that are between L_{z_i} and U_{z_i} for all $i=1,2,\ldots,K$. Thus, the combined test results from all biomarkers are $t_i=\sum_{j=1}^K\alpha_j z_{j,i}$, where $i=1,\ldots,r_T$ and $r_T=n_{rr\ldots r}$. Let $t_{(i)}$ be the ith value among t_i, thus $-\infty<\sum_{j=1}^K\alpha_j L_{z_j}<t_{(1)}<\ldots<t_{(r_T)}<\sum_{j=1}^K\alpha_j U_{z_j}<\infty$. The probability mass specifications for $T_{n+1}=\sum_{j=1}^K\alpha_j Z_{j,n+1}$ are given by the following M-function values,

$$M_{T_{n+1}}(t_{(i)},t_{(i+1)})=\frac{1}{n+1}\ \ ,\ \ i=0,1,\ldots,r_T,$$

where $t_{(0)}=-\infty$ and $t_{(r_T+1)}=\infty$. The remaining (3^K-1) M-function values can be defined as in Definition 1, e.g.

$$M_{T_{n+1}}(Z_{1,n+1}\in(-\infty,L_{z_1}),Z_{2,n+1}\in(-\infty,L_{z_2}),\ldots,Z_{K,n+1}\in(-\infty,L_{z_K}))=\frac{n_{ll\ldots l}}{n+1},$$

where $n_{ll\ldots l}$ is the number of individuals whose biomarker measurements $Z_1,Z_2,\ldots,Z_K$ are less than $L_{z_1},L_{z_2},\ldots,L_{z_K}$, respectively, and so on. We will illustrate the proposed method for combining three and four biomarkers, as introduced above, in Example 2. To this end, the data structures for combining three and four biomarkers are given in Appendix A2.

6 Evaluation

In this section, a simulation study is conducted to illustrate the proposed method for different scenarios. We have simulated $(X, Y)$from the bivariate normal distribution, for n_1 cases (disease) and n_0 controls (non-disease), with mean and variance-covariance matrix for the cases (disease) and for the controls (non-disease), respectively,

$$\mu_1 = \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \Sigma_1 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}, \mu_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \Sigma_0 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}.$$

without loss of generality we assume that $\mu_x > \mu_y > 0$, and we considered $\rho \geq 0$ to be of most practical interest [1]. The ROC curve for biomarker X alone is equal to $\{\text{ROC}_x(t) = \Phi\left(\mu_x + \Phi^{-1}(t)\right); t \in (0,1)\}$ with $\text{AUC}_x = \Phi(\mu_x/\sqrt{2})$, The means of biomarker X measurements corresponding to $\text{AUC}_x = 0.6, 0.7, 0.8, 0.9$ are $0.358, 0.742, 1.190, 1.812$, respectively. The same can be defined for biomarker Y. According to Su and Liu [6], the area under the ROC curve associated with $\alpha_1 X + \alpha_2 Y$ is $\text{AUC}^* = \Phi\left(\sqrt{(\mu_1 - \mu_0)^T(\Sigma_0 + \Sigma_1)^{-1}(\mu_1 - \mu_0)}\right)$, which is optimized at $\underline{\alpha}^*_{\mathbf{opt}} = (\Sigma_0 + \Sigma_1)^{-1}(\mu_1 - \mu_0)$, where Φ denotes the standard normal cumulative distribution function. For our setting, the area under the ROC curve associated with $\alpha_1 X + \alpha_2 Y$, with $\alpha_1, \alpha_2 \in [0,1]$ and $\alpha_1 + \alpha_2 = 1$, can be written as

$$\text{AUC} = \Phi\left(\sqrt{(\mu_x + \mu_y)(\mu_x \alpha_1 + \mu_y \alpha_2)/(2 + 2\rho)}\right). \tag{14}$$

which is optimized at

$$\underline{\alpha}_{\mathbf{opt}} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \frac{1}{(\mu_x - \rho\mu_y)^2 + (\mu_y - \rho\mu_x)^2} \begin{bmatrix} (\mu_x - \rho\mu_y)^2 \\ (\mu_y - \rho\mu_x)^2 \end{bmatrix}. \tag{15}$$

When X and Y are equally accurate on their own, i.e. $\mu_x = \mu_y = \mu$, the optimal linear combination is the average of these two biomarkers, that is $0.5X + 0.5Y$, otherwise the more accurate biomarker will be given more weight in the optimal linear combination. The area under the curve in this case is $\text{AUC} = \text{AUC}^* = \Phi\left(\mu/\sqrt{1+\rho}\right)$. So when two biomarkers of the same accuracy are combined, regardless whether they are correlated or not, the values of AUC with and without restriction on the optimal $\underline{\alpha}$ (i.e. $\underline{\alpha}_{\mathbf{opt}}$ and $\underline{\alpha}^*_{\mathbf{opt}}$) are equal.

6.1 Restriction on the optimal $\underline{\alpha}$

Before we run our simulation study, we need to discuss the convenience of using this restriction on optimal $\underline{\alpha}$, as in (14) and (15). The proposed NPI method can be used without any restriction on the optimal $\underline{\alpha}$ as

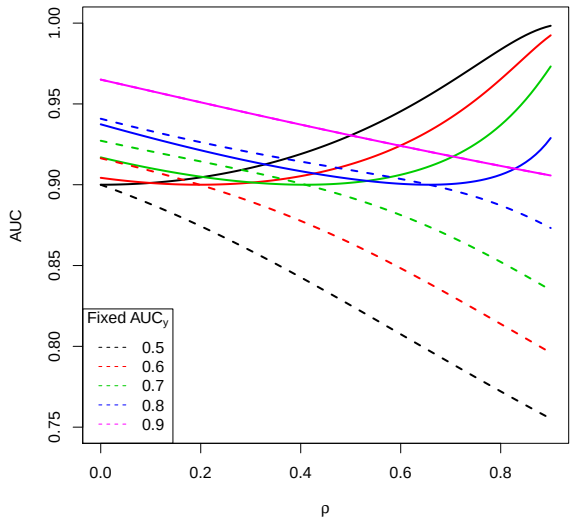

Figure 2: Fixed AUC$_y$

in [1], as will be illustrated in Example 1. In this case the complexity of the calculations for the proposed NPI method will be similar as for the method proposed by Pepe and Thompson [1]. We discuss below how the restriction on the optimal $\underline{\alpha}$ can reduce computational complexity and helps in deciding when combining diagnostic tests does actually lead to improved accuracy.

First, with the restriction placed on the optimal $\underline{\alpha}$, we only search for the optimal $\underline{\alpha} = (\alpha_1, \alpha_2)$ to be between 0 and 1 such that $\alpha_1 + \alpha_2 = 1$, which is faster than the empirical search over all possible values as discussed in [1]; further discussion is included at the end of Section 4.

In Figure 2 we have plotted the combined AUC for different values of AUC$_y$ with AUC$_x = 0.9$, over all possible values of $\rho \geq 0$. The combined AUCs with restriction are represented by dashed lines while the combined AUCs without restriction are represented by solid lines. As discussed above, when two tests of the same accuracy (AUC$_x$ = AUC$_y = 0.9$) are combined, the AUC values with and without restriction are identical (the pink lines), where combining uncorrelated tests leads to more improvement in accuracy than combining correlated tests (AUC $= 0.965$ for uncorrelated tests and AUC $= 0.9$ for correlated tests). Note that there is of course no accuracy improvement by combining two perfectly correlated tests of the same accuracy.

For other cases of combining two tests of different accuracy (AUC$_x = 0.9, 0.5 \leq$ AUC$_y \leq 0.9$), there is clear difference between the combined AUC with and without restriction. For example, the combined AUCs without restriction have a U shape, meaning that the improvement in accuracy is high by combining two uncorrelated tests, then it drops for increasing level of correlation, and then the improvement in accuracy

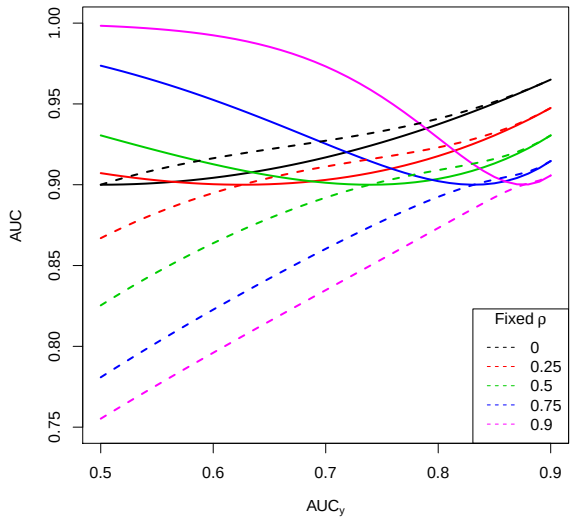

Figure 3: Fixed ρ

picks up again and reaches peak values when the two tests are strongly correlated; more discussion about this case can be found in Bansal and Pepe [29].

For the combined AUC with restriction on the optimal $\underline{\alpha}$, it holds that the more correlated the tests are the less improvement in accuracy is achieved by combining them. In other words, the best improvement in accuracy is obtained when two uncorrelated tests are combined. For example, if one combines a good test, say with $\text{AUC}_x = 0.9$, with a useless test, with $\text{AUC}_y = 0.5$, then the combined AUC without restriction takes the values $\text{AUC}^* = 0.9$ if the two tests are uncorrelated and the value $\text{AUC}^* = 1$ if the tests are perfectly correlated, which is puzzling as if these tests are really correlated why one of them behaves so good and the second behaves so badly. On the other hand, with the combined AUC with restriction, it takes the values $\text{AUC} = 0.9$ if the tests combined are uncorrelated and $\text{AUC} = 0.739$ if they are highly correlated, meaning that we are worse off by combining the two tests than considering the good test alone.

Another interesting point about using the combined AUC with restriction, from Figure 2, is that the rate of reduction in improvement depends on the accuracy of the Y test (AUC_y), that is if one combines a good test, say $\text{AUC}_x = 0.9$, with an average test, say $\text{AUC}_y = 0.8$, then the reduction in improvement if these two tests are highly correlated is smaller than the reduction in the improvement if one combines two highly correlated tests, one of which is good with $\text{AUC}_x = 0.9$ while the second is not that good, say with $\text{AUC}_y = 0.6$. Furthermore, for uncorrelated or weakly correlated tests, the improvement gained by combining tests, using the combined AUC with restriction, is equal or greater than for the combined AUC without restriction, meaning that we can find the optimal combination of uncorrelated or weakly correlated

tests with much higher improvement by using the combined AUC with restriction.

In Figure 3 we have plotted the combined AUC for different values of ρ over all possible values of $0.5 \leq \text{AUC}_y \leq 0.9$, where $\text{AUC}_x = 0.9$. The combined AUCs with restriction are represented by dashed lines and the combined AUCs without restriction by solid lines. Figure 3 shows the same behaviours as mentioned above, that is the combined AUC without restriction is U-shaped while the combined AUC with restriction is monotone. Both combined AUCs agree when two tests of the same accuracy are combined, where higher values of the combined AUCs are associated with uncorrelated tests. The combined AUCs with and without restriction differ significantly when one combines tests of different accuracy. For fixed ρ, the greater AUC_y is, the greater improvement is achieved by the combined AUC with restriction, while this nice property does not hold for the combined AUC without the restriction. For example, for $\rho = 0.5$ and $\text{AUC}_x = 0.9$, the green lines in Figure 3, when $\text{AUC}_y = 0.50, 0.75, 0.90$ the combined AUC without restriction is 0.931, 0.900, 0.931, respectively, while the corresponding combined AUC with restriction is 0.825, 0.902, 0.931, respectively.

For uncorrelated tests, the combined AUC with restriction leads to overall greater improvement (higher AUC) than the combined AUC without restriction. For the combined AUC with restriction, the rate of improvement of combining uncorrelated tests is equal to the accuracy of test X (that is AUC_x) if we combined this test with a useless test ($\text{AUC}_y = 0.5$), so basically the weight given to the useless test is zero. The maximum value of $\Phi(\sqrt{(}2)\Phi^{-1}(AUC_x)) = \Phi(\mu_x)$ is obtained if two unrelated tests of the same accuracy are combined ($\text{AUC}_x = \text{AUC}_y$), in this case the optimal weights are (0.5,0.5).

6.2 Simulation study

The results of the simulation study are based on 1000 simulations for $n_1 = n_0 = 50, 100$, $\rho = 0, 0.5, 0.75$, and for different values of μ_x and μ_y. First we have considered the case when there is no LoD, i.e. complete data, then we have introduced some LoD per group, such as L and U to be equal to the 10th and 90th quantiles of the simulated sample, respectively. The results of this simulation are given in Table 5 for the case 'without LoD', and in Table 6 for the case 'with LoD'.

From Table 5, for the case 'without LoD', we notice that NPI lower and upper AUCs always bound the empirical one, this is because NPI approach is exactly calibrated [30, 31], in the sense that it never leads to results that are in conflict with inferences based on empirical probabilities, in our case the empirical AUC. We also notice that the empirical AUC is much closer to the upper AUC than to the lower AUC, this may correspond to the fact the AUC is over optimistic when considering the predictive performance [5]. These two remarks highlight the merits of the proposed NPI methods, e.g. for the latter it captured the over

optimistic behaviour of the empirical AUC. We cannot, however, generalize this to the case 'with LoD', in Table 6, the main reason is that the empirical AUC is calculated after removing the observations that are subject to LoD, while in NPI method the number of these observations is taken into account to derive the NPI lower and upper AUC.

For the case without LoD, Table 5, the values of the optimal $\underline{\alpha}$ returned by the NPI method, are close or sometimes identical to the optimal $\underline{\alpha}$ returned by the empirical AUC, this is again due to the calibrated property of the NPI. However, this is not the case for the inference with LoD, Table 6, as the values of the optimal $\underline{\alpha}$ returned by the NPI method are different from the empirical one, yet the optimal $\underline{\alpha}$ corresponding to the upper AUC is much closer to the empirical one than the one corresponding to the lower AUC. These results are due to the same reasons discussed above (that is how the AUC is calculated, and the fact that empirical AUC is over optimistic).

From Table 5, when two biomarkers of the same accuracy (the same AUC) are combined, the empirical AUC and the lower and upper NPI AUCs all give equal weights to both biomarkers, that is $\underline{\alpha}_{opt} = (0.5, 0.5)$. It is also interesting to see that if two uncorrelated tests with the same values of AUC are combined then that leads to more improvement than combining two correlated tests of the same accuracy. In fact, this holds regardless whether the two tests have the same accuracy or not, that is for the case 'without LoD', combining two uncorrelated tests can lead to more improvement than combining two correlated tests. This holds for the empirical AUC and the NPI lower and upper AUC. While the same holds for the case with LoD, Table 6, for both the empirical AUC and upper AUC, it does not hold for the lower AUC. Considering the lower AUC, one can achieve higher improvement by combining two correlated tests, this is due to the fact that now we have less information in favour of the event of interest.

To study the robustness of the NPI proposed method, a simulation study from an asymmetric distribution is performed. In our context the Gamma distribution is often used. As the NPI method is nonparametric, in the sense that we do not make any assumption about the underlying distribution, it is expected that its performance will not be affected by the choice of the underlying distribution.

Considering a biomarker X, the area under the ROC curve that is constructed from the Gamma distributions of the disease and non-disease groups with shape and scale parameters γ_x^1, β_x^1 and γ_x^0, β_x^0, respectively, is given by [8]

$$AUC_x = \Psi_x \left(\frac{\beta_x^1}{\beta_x^1 + \beta_x^0} \right),$$

where Ψ_x is the Beta cumulative distribution function with parameters (γ_x^0, γ_x^1). Similarly, one can define the area under the ROC curve for biomarker Y, AUC_y, that is constructed from the Gamma distributions

of the disease and non-disease groups with shape and scale parameters γ_y^1, β_y^1 and γ_y^0, β_y^0, respectively, as

$$AUC_y = \Psi_y \left(\frac{\beta_y^1}{\beta_y^1 + \beta_y^0} \right),$$

where Ψ_y is the Beta cumulative distribution function with parameters (γ_y^0, γ_y^1).

We have simulated (X, Y) from the Bigamma distribution, for n_1 cases (disease) and n_0 controls (non-disease), with shape and scale parameters set to $\gamma_x^1 = 1$, $\beta_x^1 = 1$ and $\gamma_y^1 = 1$, $\beta_y^1 = 1$ for the disease group. For the non-disease group the shape parameters are set to ($\gamma_x^0 = 1$, $\gamma_y^0 = 1$), and the scale parameters (β_x^0, β_y^0) are set to achieve $AUC_x = 0.6, 0.7, 0.8, 0.9$, and $AUC_y = 0.6, 0.7, 0.8, 0.9$, respectively. Following [32], the Bigamma distribution was constructed using the Gaussian copula with correlation coefficient $\rho = 0, 0.5, 0.75$ and with the above specified Gamma marginal distributions.

The simulation results from the Bigamma distribution are summarized in Tables 7 and 8 in the appendix. From these tables we observed the same results as the Normal distribution case which demonstrates the robustness of the proposed NPI method.

7 Examples

Example 1 (**Pancreatic cancer data set).** In this example, we use the data set from a study of 90 pancreatic cancer patients and 51 control patients with pancreatitis [33]. Two serum markers were measured on these patients, the cancer antigen CA125 and the carbohydrate antigen CA19-9. The marker values were transformed to a natural logarithmic scale and are displayed in Figure 4. For ease of presentation, let `log(CA19-9)` be the biomarker X and `log(CA125)` the biomarker Y. Three scenarios are considered below. In scenario A, we consider the whole data set, i.e. without LoD. In scenario B, we introduce the following LoD scheme, $L_x = 4.5$, $U_x = 5000$, $L_y = 6$ and $U_y = 100$, where ($l_x^0 = 5, r_x^0 = 46, u_x^0 = 0$), ($l_x^1 = 4, r_x^1 = 77, u_x^1 = 9$), ($l_y^0 = 4, r_y^0 = 45, u_y^0 = 2$) and ($l_y^1 = 1, r_y^1 = 82, u_y^1 = 7$). In scenario C, we have the LoD scheme $L_x = 11$, $U_x = 5000$, $L_y = 6$ and $U_y = 80$, where ($l_x^0 = 28, r_x^0 = 23, u_x^0 = 0$), ($l_x^1 = 11, r_x^1 = 70, u_x^1 = 9$), ($l_y^0 = 4, r_y^0 = 44, u_y^0 = 3$) and ($l_y^1 = 1, r_y^1 = 78, u_y^1 = 11$). To make sure that our biomarkers tests results are comparable we use the standardized values (i.e. with mean 0 and variance 1) after the natural logarithmic transformation. The objective is to find the best linear combination of X and Y which yields higher AUC value than either one alone.

Under these three scenarios, the values of α (α_1 and α_2) that maximize the empirical AUC and the NPI lower and upper AUC are given in Table 1. For the sake of comparison we also obtained the empirical AUC and the NPI lower and upper AUC for each biomarker. For scenarios B and C, the empirical AUC values are calculated after discarding (removing) the observations that are below or above the LoD for each biomarker.

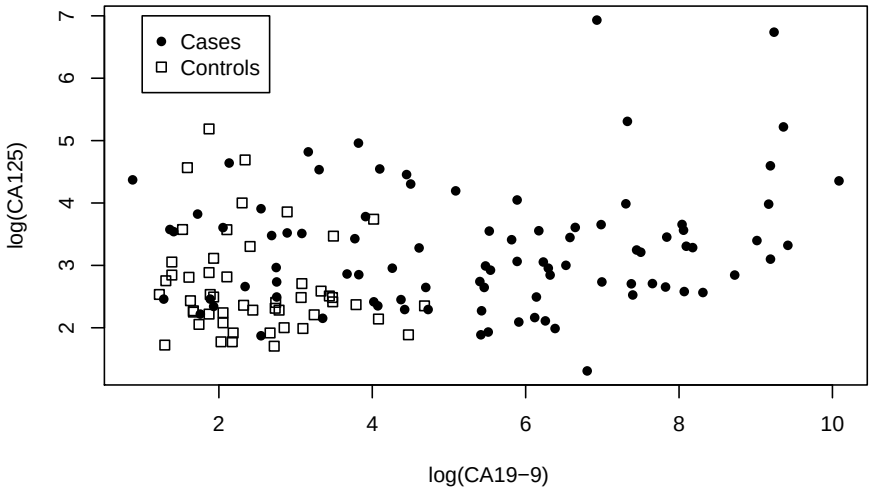

Figure 4: Pancreatic cancer data set (Example 1)

Scenario	Biomarkers	$\hat{\alpha}_{opt}$	$\widehat{AUC}$	α^L_{opt}	$\underline{AUC}$	α^U_{opt}	$\overline{AUC}$
A	X		0.8614		0.8347		0.8664
	Y		0.7056		0.6830		0.7158
	$X + \alpha Y$	(1, 0.39)	0.8937	(1, 0.39)	0.8669	(1, 0.39)	0.8969
	$\alpha_1 X + \alpha_2 Y$	(0.720, 0.280)	0.8937	(0.720, 0.280)	0.8669	(0.720, 0.280)	0.8969
B	X		0.8755		0.8328		0.8688
	Y		0.6931		0.6809		0.7175
	$X + \alpha Y$	(1, 0.39)	0.8939	(1, 0.42)	0.7494	(1, 0.42)	0.9216
	$\alpha_1 X + \alpha_2 Y$	(0.724, 0.276)	0.8939	(0.705, 0.295)	0.7496	(0.705, 0.295)	0.9218
C	X		0.8696		0.8068		0.9024
	Y		0.6936		0.6792		0.7198
	$X + \alpha Y$	(1, 0.39)	0.9069	(1, 0.07)	0.6796	(1, 0.48)	0.9704
	$\alpha_1 X + \alpha_2 Y$	(0.724, 0.276)	0.9069	(0.931, 0.069)	0.6813	(0.675, 0.325)	0.9704

Table 1: Pancreatic cancer data set results (Example 1)

For the NPI lower and upper AUC, these observations are not removed but only their numbers are taken into account. This is the reason why the empirical AUC is no longer always between the NPI lower and upper AUCs. The data structures of these two scenarios are given below, for example for scenario B the results are based on 71 cases and 41 controls while for scenario C the results are based on 62 cases and 22 controls.

$$\mathbb{S}_B^{D=0} = \begin{bmatrix} 0 & 0 & 0 \\ 3 & 41 & 2 \\ 1 & 4 & 0 \end{bmatrix}, \mathbb{S}_B^{D=1} = \begin{bmatrix} 0 & 7 & 2 \\ 1 & 71 & 5 \\ 0 & 4 & 0 \end{bmatrix}, \mathbb{S}_C^{D=0} = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 22 & 0 \\ 3 & 22 & 3 \end{bmatrix}, \mathbb{S}_C^{D=1} = \begin{bmatrix} 0 & 6 & 3 \\ 1 & 62 & 7 \\ 0 & 10 & 1 \end{bmatrix}.$$

From Table 1, we can see that maximizing the AUC by finding the values (α_1, α_2) in $\alpha_1 X + \alpha_2 Y$, where $\alpha_1 + \alpha_2 = 1$, gives slightly higher value for AUC compared to maximizing the AUC by finding the value

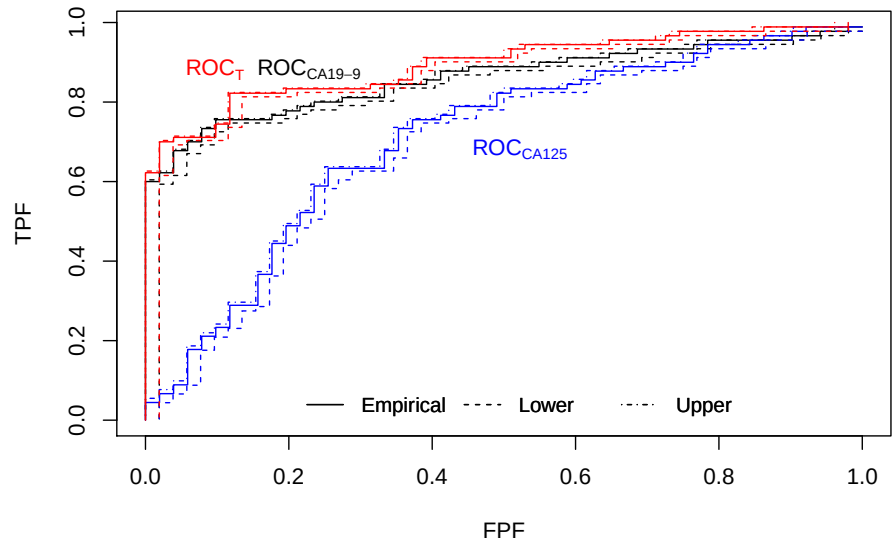

Figure 5: Lower, empirical and upper ROC curves for scenario A (Example 1)

α in $X + \alpha Y$. For scenarios A and B, CA19-9 has been given around 70% weight while CA125 has been given a weight of about 30%. For scenario C, the result is quite different, with the lower AUC maximized by assigning higher weight to CA19-9 compared to CA125, while for the optimal upper AUC more weight is given to CA125. The optimal values for the empirical, lower and upper AUCs are close to each other in scenario A, i.e. without LoD.

For scenario A, combining two biomarkers leads to accuracy improvement, that is larger values of (empirical, lower and upper) AUC compared to the AUC values of the individual ones, this is illustrated in Figure 5, where $T = 0.72X + 0.28Y$. This improvement is quite small in comparison to using CA19-9 alone, this may be due to the fact that we have combined a good biomarker (CA19-9) with an average biomarker (CA125).

For scenarios B and C (both with LoD) we notice that the lower AUC for combining two biomarkers is not greater than the individual lower AUC values, while the upper AUC for combining two biomarkers is much greater than the individual upper AUC values, in fact it is much greater than the upper AUC of combining two biomarkers for scenario A. This is because with LoD we have fewer observations for which both biomarkers results are available, that is we have less evidence in favour (against) of the event of interest (AUC here) for the lower (upper) AUC. Furthermore, the imprecision is larger for scenarios B and C compared to scenario A, as with LoD we have fewer observations in which the biomarkers results are both available.

Example 2 (DMD data set). The data set used in this example results from a study to develop screening

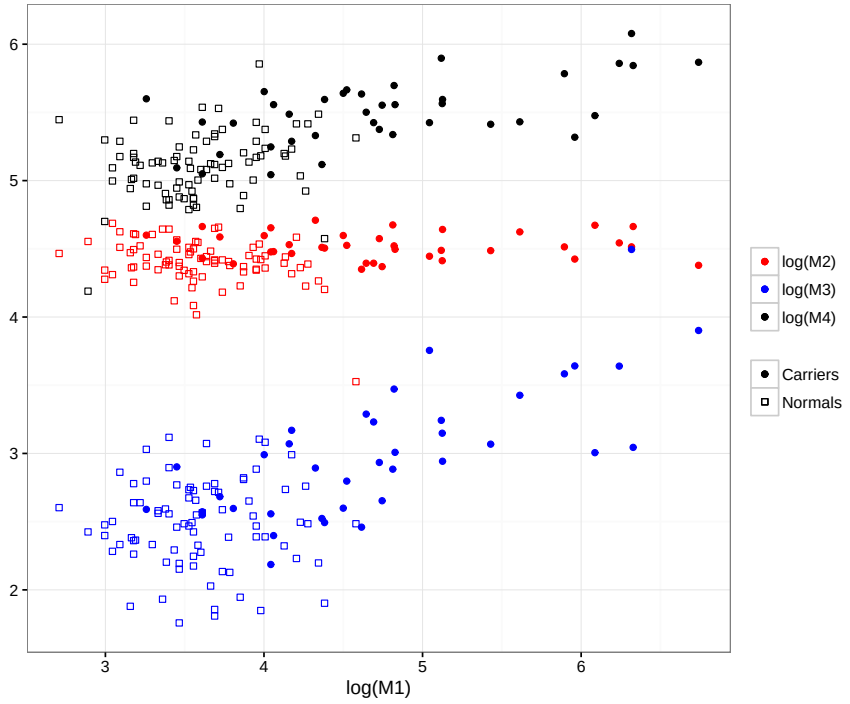

Figure 6: DMD data set (Example 2)

methods to identify carriers of a rare genetic disorder. Four measurements M1, M2, M3 and M4 were made on blood samples. The data were first discussed by Cox *et al.* [34] and are available from Carnegie Mellon University Statlib Datasets Archive at `ftp://rcom.univie.ac.at/mirrors/lib.stat.cmu.edu/datasets/.index.html`. There are several samples for some patients, for which the averages are considered, and five missing values are excluded from the analysis. The remaining sample, which is used in this example, consists of 120 observations, 82 'normals' and 38 'carriers'. The four measurements were transformed to a natural logarithmic scale and are displayed in Figure 6. As in the previous example, we use the standardized values after the natural logarithmic transformation in the analysis. From the correlation matrix given below, we see that M1 is quite strongly correlated with M3 and M4, and M3 is quite strongly correlated with M4, yet M2 is only weakly correlated with the other measurements.

$$\text{Corr} = \begin{array}{c} \\ M1 \\ M2 \\ M3 \\ M4 \end{array} \begin{array}{c} \begin{array}{cccc} M1 & M2 & M3 & M4 \end{array} \\ \begin{pmatrix} 1.000 & 0.115 & 0.644 & 0.642 \\ 0.115 & 1.000 & 0.221 & 0.284 \\ 0.644 & 0.221 & 1.000 & 0.561 \\ 0.642 & 0.284 & 0.561 & 1.000 \end{pmatrix} \end{array}.$$

In this example we introduce the limits of detection as presented in Table 2, that is considering M1 with

					without LoD			with LoD		
Measurements	L	U	(l^0, r^0, u^0)	(l^1, r^1, u^1)	$\widehat{AUC}$	$\underline{AUC}$	$\overline{AUC}$	$\widehat{AUC}$	$\underline{AUC}$	$\overline{AUC}$
M1	20.5	400	(4,78,0)	(0,33,5)	0.9034	0.8684	0.9082	0.8831	0.8684	0.9082
M2	66	106.5	(5,76,1)	(0,35,3)	0.7526	0.7223	0.7640	0.7241	0.7220	0.7646
M3	6.8	37	(6,76,0)	(0,33,5)	0.8232	0.7912	0.8310	0.7803	0.7912	0.8310
M4	115	347	(3,78,1)	(0,34,4)	0.8789	0.8446	0.8848	0.8705	0.8434	0.8848

Table 2: DMD data set (Example 2)

LoD, we have 111 observations within LoD, 78 normals and 33 carriers. In this table we have calculated the empirical, lower and upper AUC for the individual biomarkers, with and without LoD. For the case with LoD, the empirical AUC values are calculated after discarding (removing) the observations that are below or above the LoD for each biomarker. Obviously, the NPI lower and upper AUC values bound the empirical AUC for the case without LoD, while this is not necessary for the case with LoD as we have discussed before. From this table we notice that M1 has the largest accuracy (AUC), then M4 and M3, and M2 has the smallest accuracy. From NPI perspective, as $\underline{\text{AUC}}_{M1} > \overline{\text{AUC}}_{M3} > \overline{\text{AUC}}_{M2}$ we can say that there is a strong indication that M1 is better than M2 and M3, and as we have $\underline{\text{AUC}}_{M1} > \underline{\text{AUC}}_{M4}$ and $\overline{\text{AUC}}_{M1} > \overline{\text{AUC}}_{M4}$ we can say that there is a weak indication that M1 is better than M4, for more details about using NPI for pairwise and multiple comparisons we refer to [35].

Measurements	$\hat{\alpha}_{opt}$	$\widehat{AUC}$	α^L_{opt}	$\underline{AUC}$	α^U_{opt}	$\overline{AUC}$
			Without LoD			
$\alpha_1 M_1 + \alpha_2 M_2$	(0.609 , 0.391)	0.9535	(0.609 , 0.391)	0.9178	(0.609 , 0.391)	0.9552
$\alpha_1 M_1 + \alpha_2 M_3$	(0.750 , 0.250)	0.9178	(0.750 , 0.250)	0.8835	(0.750 , 0.250)	0.9209
$\alpha_1 M_1 + \alpha_2 M_4$	(0.551 , 0.449)	0.9313	(0.551 , 0.449)	0.8965	(0.551 , 0.449)	0.9339
$\alpha_1 M_2 + \alpha_2 M_3$	(0.500 , 0.500)	0.8780	(0.500 , 0.500)	0.8452	(0.500 , 0.500)	0.8826
$\alpha_1 M_2 + \alpha_2 M_4$	(0.306 , 0.694)	0.9095	(0.306 , 0.694)	0.8755	(0.306 , 0.694)	0.9129
$\alpha_1 M_3 + \alpha_2 M_4$	(0.256 , 0.744)	0.9156	(0.256 , 0.744)	0.8814	(0.256 , 0.744)	0.9188
			With LoD			
$\alpha_1 M_1 + \alpha_2 M_2$	(0.609 , 0.391)	0.9368	(0.499 , 0.501)	0.8576	(0.609 , 0.391)	0.9558
$\alpha_1 M_1 + \alpha_2 M_3$	(0.879 , 0.121)	0.8898	(0.581 , 0.419)	0.8171	(0.879 , 0.121)	0.9240
$\alpha_1 M_1 + \alpha_2 M_4$	(0.520 , 0.480)	0.9211	(0.520 , 0.480)	0.8783	(0.520 , 0.480)	0.9376
$\alpha_1 M_2 + \alpha_2 M_3$	(0.580 , 0.420)	0.8248	(0.460 , 0.540)	0.7519	(0.500 , 0.500)	0.8845
$\alpha_1 M_2 + \alpha_2 M_4$	(0.306 , 0.694)	0.8943	(0.526 , 0.474)	0.8026	(0.340 , 0.660)	0.9203
$\alpha_1 M_3 + \alpha_2 M_4$	(0.256 , 0.744)	0.9024	(0.493 , 0.507)	0.8455	(0.416 , 0.584)	0.9237

Table 3: DMD data set, two measurements are combined (Example 2)

In Table 3, we combine two measurements (without and with LoD, resp.) in order to maximize the areas under the empirical, lower and upper ROC curves. From Table 3, without LoD, it seems that combining M1 with M2 gives the largest improvement in comparison to combining M1 with M3 or M1 with M4, despite the fact that M3 and M4 have the highest AUC alone compared to M2. However M3 and M4 are highly correlated with M1 while M2 is uncorrelated with M1. The same holds for the results in Table 3, with LoD, for both the empirical AUC and upper AUC, but for the lower AUC we have different results. For the lower AUC, the largest improvement is obtained by combining M1 and M4, then the second largest by combining

M1 with M2, this is again due to the fact that M1 and M4 have higher AUC alone and they are highly correlated, but also to the fact that we have fewer observations with LoD, as discussed in Section 6. The data structures for the analysis in Table 3 with LoD, are given below, e.g. considering M1 with M2, the results are based on 103 observations within LoD, 72 normals and 31 carriers.

$$\mathbb{S}_{12}^{D=0}=\begin{bmatrix}0&0&0\\5&72&1\\0&4&0\end{bmatrix},\mathbb{S}_{12}^{D=1}=\begin{bmatrix}0&4&1\\0&31&2\\0&0&0\end{bmatrix},\mathbb{S}_{13}^{D=0}=\begin{bmatrix}0&0&0\\6&72&0\\0&4&0\end{bmatrix},\mathbb{S}_{13}^{D=1}=\begin{bmatrix}0&2&3\\0&31&2\\0&0&0\end{bmatrix},$$

$$\mathbb{S}_{14}^{D=0}=\begin{bmatrix}0&0&0\\1&76&1\\2&2&0\end{bmatrix},\mathbb{S}_{14}^{D=1}=\begin{bmatrix}0&2&3\\0&32&1\\0&0&0\end{bmatrix},\mathbb{S}_{23}^{D=0}=\begin{bmatrix}0&1&0\\6&70&0\\0&5&0\end{bmatrix},\mathbb{S}_{23}^{D=1}=\begin{bmatrix}0&3&0\\0&30&5\\0&0&0\end{bmatrix},$$

$$\mathbb{S}_{24}^{D=0}=\begin{bmatrix}0&1&0\\3&72&1\\0&5&0\end{bmatrix},\mathbb{S}_{24}^{D=1}=\begin{bmatrix}0&3&0\\0&31&4\\0&0&0\end{bmatrix},\mathbb{S}_{34}^{D=0}=\begin{bmatrix}0&0&0\\2&73&1\\1&5&0\end{bmatrix},\mathbb{S}_{34}^{D=1}=\begin{bmatrix}0&2&3\\0&32&1\\0&0&0\end{bmatrix}.$$

In Table 4, we combine three measurements (without and with LoD, respectively) in order to maximize the areas under the empirical, lower and upper ROC curves. For combining three and four markers we have used the approach discussed at the end of Section 5. By comparing Tables 3 and 4, the case without LoD, one can see that we gain little improvement by combining M3 or M4 with M1 and M2. On the other hand, combining M1 with M3 and M4 leads to better improvement compared to combining M2 with M3 and M4. The empirical, lower and upper AUCs all agreed on this.

Measurements	$\hat{\alpha}_{opt}$	$\widehat{AUC}$	α_{opt}^{L}	$\underline{AUC}$	α_{opt}^{U}	$\overline{AUC}$
			Without LoD			
$\alpha_1M_1+\alpha_2M_2+\alpha_3M_3$	(0.52, 0.34, 0.14)	0.9589	(0.52, 0.34, 0.14)	0.9231	(0.52, 0.34, 0.14)	0.9605
$\alpha_1M_1+\alpha_2M_2+\alpha_3M_4$	(0.54, 0.32, 0.14)	0.9589	(0.54, 0.32, 0.14)	0.9231	(0.54, 0.32, 0.14)	0.9605
$\alpha_1M_1+\alpha_2M_3+\alpha_3M_4$	(0.31, 0.19, 0.50)	0.9336	(0.31, 0.19, 0.50)	0.8987	(0.31, 0.19, 0.50)	0.9361
$\alpha_1M_2+\alpha_2M_3+\alpha_3M_4$	(0.27, 0.28, 0.45)	0.9272	(0.27, 0.28, 0.45)	0.8925	(0.27, 0.28, 0.45)	0.9299
			With LoD			
$\alpha_1M_1+\alpha_2M_2+\alpha_3M_3$	(0.52, 0.34, 0.14)	0.9347	(0.50, 0.50, 0)	0.7402	(0.52, 0.34, 0.14)	0.9614
$\alpha_1M_1+\alpha_2M_2+\alpha_3M_4$	(0.51, 0.27, 0.22)	0.9471	(0.49, 0.50, 0.01)	0.8128	(0.51, 0.27, 0.22)	0.9657
$\alpha_1M_1+\alpha_2M_3+\alpha_3M_4$	(0.31, 0.19, 0.50)	0.9183	(0.47, 0.28, 0.25)	0.7943	(0.31, 0.19, 0.50)	0.9435
$\alpha_1M_2+\alpha_2M_3+\alpha_3M_4$	(0.29, 0.21, 0.50)	0.9048	(0.04, 0.49, 0.47)	0.7115	(0.27, 0.28, 0.45)	0.9416

Table 4: DMD data set, three measurements are combined (Example 2)

For the case with LoD, by comparing Tables 3 and 4, we gain more accuracy by combining M4 with M1 and M2 than with M3, for both the empirical and the upper AUC, and a small loss of accuracy for the lower AUC (again the lower AUC is affected by having fewer observations). It does not make a big difference if we combine M1 or M2 with M3 and M4 for both the empirical and the upper AUC, but there is a large loss in the accuracy in terms of the lower AUC if we combine M2 instead of M1 with M3 and M4. The data

structures for the analysis in Table 3 with LoD, are given below, e.g. considering M1, M2 and M3, the results is based on 95 observations, 66 normals and 29 carriers.

$$\mathbb{S}_{123}^{D=0}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\\0&1&0\\6&66&0\\0&5&0\\0&0&0\\0&4&0\\0&0&0\end{bmatrix},\mathbb{S}_{123}^{D=1}=\begin{bmatrix}0&1&0\\0&1&3\\0&0&0\\0&2&0\\0&29&2\\0&0&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix},\mathbb{S}_{124}^{D=0}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\\0&1&0\\1&70&1\\0&5&0\\0&0&0\\2&2&0\\0&0&0\end{bmatrix},\mathbb{S}_{124}^{D=1}=\begin{bmatrix}0&1&0\\0&1&3\\0&0&0\\0&2&0\\0&30&1\\0&0&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix},$$

$$\mathbb{S}_{134}^{D=0}=\begin{bmatrix}0&0&0\\0&0&0\\0&0&0\\0&0&0\\0&71&1\\1&5&0\\0&0&0\\2&2&0\\0&0&0\end{bmatrix},\mathbb{S}_{134}^{D=1}=\begin{bmatrix}0&0&3\\0&2&0\\0&0&0\\0&2&0\\0&30&1\\0&0&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix},\mathbb{S}_{234}^{D=0}=\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\\0&0&0\\2&67&1\\1&5&0\\0&0&0\\0&5&0\\0&0&0\end{bmatrix},\mathbb{S}_{234}^{D=1}=\begin{bmatrix}0&0&0\\0&3&0\\0&0&0\\0&2&3\\0&29&1\\0&0&0\\0&0&0\\0&0&0\\0&0&0\end{bmatrix}.$$

Finally, if we combine all 4 measurements, in the case without LoD, we have $\hat{\alpha}_{opt}=\alpha_{opt}^{L}=\alpha_{opt}^{U}=(0.52,0.31,0.08,0.09)$ where $\widehat{AUC}=0.9621$, $\underline{AUC}=0.9262$ and $\overline{AUC}=0.9635$. That is about 83% of the weight is given to the first and the second measurements (52% for M1 and 31% for M2), and M3 and M4 are almost neglected. For the case with LoD (as shown from the data structure given below, the result is based on 93 observations, 65 normals and 28 carriers) we have $\hat{\alpha}_{opt}=(0.45,0.27,0.04,0.24)$ and $\widehat{AUC}=0.9451$, $\alpha_{opt}^{L}=(0.49,0.5,0,0.01)$ and $\underline{AUC}=0.7105$, $\alpha_{opt}^{U}=(0.45,0.27,0.04,0.24)$ and $\overline{AUC}=0.9691$. In this case, in order to maximize the NPI lower AUC, M1 and M2 are given almost the same weights (around 50% each) while M3 and M4 are almost totally neglected. The empirical and the upper AUC still gave more weight to M1 (45%) but they also assigned M2 and M4 almost similar weights (27% and 24%, resp.), while M3 is still neglected.

$$\mathbb{S}_{1234}^{D=0} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 2 & 2 & 0 & 0 & 65 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}, \mathbb{S}_{1234}^{D=1} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 & 28 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

One can see from Table 4, the case without LoD, and the results above, that if we add M4, the empirical, and upper AUCs become larger, so we improved the accuracy by combining all markers together. On the other hand, for the case with LoD, we observed a similar behaviour as in Example 1, namely that the empirical and the upper AUC improved by combining all markers together, but the lower AUC decreases which reflects that we have less evidence in favour of the event of interest, of course due to having fewer observations.

8 Concluding remarks

In this paper the NPI approach is presented for best linear combination of biomarkers subject to limits of detection. We also showed how NPI can be used to combine biomarkers when all the data are available (i.e. without limits of detection). NPI provides an attractive approach to quantify uncertainty without discarding the unobserved measurements (that are below or above limits of detection) or the need for replacement strategies and their drawbacks [8]. In NPI only the number of the unobserved measurements are taking into account and the uncertainty is quantified via lower and upper probabilities to provide a statement about the future observations.

While there are some issues regarding classical methods with regard to when combining biomarkers will actually improve the accuracy [29], the proposed method, as shown in the simulation study, with the restrictions on the coefficients of the linear combination, can help to draw a conclusion on when combining biomarkers will improve diagnostic accuracy. The best scenario is achieved by combining good (high AUC values) uncorrelated biomarkers, and the worst scenario is obtained when a good biomarker is combined with a highly correlated useless biomarker. Therefore, from practical perspective, if one chooses the AUC as the objective function to maximize, it will be more informative and computationally efficient to use restriction on the parameters. And when selecting a set of biomarkers to combine, one should take the correlation

between these biomarkers into account as mentioned earlier.

NPI typically leads to lower and upper probabilities for events of interest, which are based on Hills assumption $A_{(n)}$ and have strong properties from frequentist statistics perspective. As events of interest are explicitly about a future observation, or a function of such an observation, NPI is indeed explicitly about prediction. The NPI lower and upper probabilities have a frequentist interpretation that could be regarded as confidence statements related to repeated application of the same procedure. From this perspective, corresponding lower and upper probabilities can be interpreted as bounds for the confidence level for the event of interest. However, the method does not provide prediction intervals in the classical sense, as e.g. appear in frequentist regression methods. Those tend to relate to confidence intervals for model parameter estimates combined with variability included in the model, in NPI no variability is explicitly included in a model and there are clearly no parameters to be estimated.

The proposed method can be extended in many ways, for example, one can consider other objective functions to optimize instead of the AUC, for example by building a risk score function [2], to accommodate other factors such as costs or risk to patients, etc. We may also want to consider other ways of combining the tests, e.g. by using copulas to capture the dependence of these biomarkers. Some initial results of using NPI for combining two biomarkers via copula have been presented by Muhammad [36]. Generalizing the proposed NPI method to three-group ROC surface and the volume under the surface [24, 37] is an interesting topic for future research. NPI lower and upper bounds for the well-known Youden index [38, 39] have been introduced for both two-group ROC and three-group ROC analysis [23, 24]. Full investigation of using NPI for selecting the optimal cut-off points in the case of limits of detection [40] is of also of interest and left for future research.

Appendix

A1. Proof of Theorem 3

In this section we derive the lower and upper probabilities for the event $T^0_{n_0+1} < T^1_{n_1+1}$, which is equivalent to finding the lower and upper probabilities for the event $\alpha_1 X^0_{n_0+1} + \alpha_2 Y^0_{n_0+1} < \alpha_1 X^1_{n_1+1} + \alpha_2 Y^1_{n_1+1}$.

$$
\begin{aligned}
P =& P(T^0_{n_0+1} < T^1_{n_1+1}) = \sum_{i=0}^{r^1_T} P(T^0_{n_0+1} < T^1_{n_1+1}, T^1_{n_1+1} \in (t^1_{(i)}, t^1_{(i+1)})) \\
&+ P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (-\infty, L_x), Y^1_{n_1+1} \in (-\infty, L_y)) + P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (-\infty, L_x), Y^1_{n_1+1} \in (L_y, U_y)) \\
&+ P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (-\infty, L_x), Y^1_{n_1+1} \in (U_y, \infty)) + P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (L_x, U_x), Y^1_{n_1+1} \in (-\infty, L_y)) \\
&+ P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (L_x, U_x), Y^1_{n_1+1} \in (U_y, \infty)) + P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (U_x, \infty), Y^1_{n_1+1} \in (-\infty, L_y)) \\
&+ P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (U_x, \infty), Y^1_{n_1+1} \in (L_y, U_y)) + P(T^0_{n_0+1} < T^1_{n_1+1}, X^1_{n_1+1} \in (U_x, \infty), Y^1_{n_1+1} \in (U_y, \infty)).
\end{aligned}
$$

The NPI lower probability for the event $T^0_{n_0+1} < T^1_{n_1+1}$ is obtained as follows:

$$
\begin{aligned}
P \geq& \sum_{i=0}^{r^1_T} P(T^0_{n_0+1} < t^1_{(i)}) \frac{1}{n_1+1} + P(T^0_{n_0+1} < -\infty) \left[\frac{n^1_{ll}}{n_1+1} + \frac{n^1_{lr}}{n_1+1} + \frac{n^1_{lu}}{n_1+1} + \frac{n^1_{rl}}{n_1+1} + \frac{n^1_{ul}}{n_1+1} \right] \\
&+ P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 U_y) \frac{n^1_{ru}}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 L_y) \frac{n^1_{ur}}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 U_y) \frac{n^1_{uu}}{n_1+1} \\
=& \sum_{i=1}^{r^1_T} P(T^0_{n_0+1} < t^1_{(i)}) \frac{1}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 U_y) \frac{n^1_{ru}}{n_1+1} \\
&+ P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 L_y) \frac{n^1_{ur}}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 U_y) \frac{n^1_{uu}}{n_1+1}.
\end{aligned}
$$

The above inequality follows by putting all probability masses for $T^1_{n_1+1}$, according to the M-functions in Definition 1, corresponding to the intervals $(t^1_{(i)}, t^1_{(i+1)})$ $(i = 0, ..., r^1_T)$, $(-\infty, \alpha_1 L_x + \alpha_2 L_y)$, $(-\infty, \alpha_1 L_x + \alpha_2 U_y)$, $(-\infty, \alpha_1 U_x + \alpha_2 L_y)$,$(\alpha_1 L_x + \alpha_2 U_y, \infty)$, $(\alpha_1 U_x + \alpha_2 L_y, \infty)$, $(\alpha_1 U_x + \alpha_2 U_y, \infty)$ and $(-\infty, \infty)$ to the left-end points of these intervals. The next step in the proof is divided into four parts, I_1, I_2, I_3 and I_4. The following inequalities (in I_1, I_2, I_3 and I_4) follow by putting all probability masses for $T^0_{n_0+1}$ corresponding to the intervals $(t^0_{(j)}, t^0_{(j+1)})$ $(j = 0, ..., r^0_T)$, $(-\infty, \alpha_1 L_x + \alpha_2 L_y)$, $(-\infty, \alpha_1 L_x + \alpha_2 U_y)$, $(-\infty, \alpha_1 U_x + \alpha_2 L_y)$,$(\alpha_1 L_x + \alpha_2 U_y, \infty)$, $(\alpha_1 U_x + \alpha_2 L_y, \infty)$, $(\alpha_1 U_x + \alpha_2 U_y, \infty)$ and $(-\infty, \infty)$ to the right-end points of these intervals.

$$
\begin{aligned}
I_1 =& \sum_{i=1}^{r_T^1} P(T_{n_0+1}^0 < t_{(i)}^1) \frac{1}{n_1+1} \\
\geq& \sum_{i=1}^{r_T^1} \frac{1}{n_1+1} \Bigg[\mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y < t_{(i)}^1\} \frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t_{(i)}^1\} \frac{n_{lr}^0}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t_{(i)}^1\} \frac{n_{rl}^0}{n_0+1} \\
&+ \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < t_{(i)}^1\} \frac{1}{n_0+1} + \mathbf{1}\{\infty < t_{(i)}^1\} \left(\frac{n_{lu}^0 + n_{ru}^0 + n_{ul}^0 + n_{ur}^0 + n_{uu}^0 + 1}{n_0+1} \right) \Bigg] \\
=& \frac{1}{(n_1+1)(n_0+1)} \sum_{i=1}^{r_T^1} \Bigg[n_{ll}^0 + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t_{(i)}^1\} n_{lr}^0 + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t_{(i)}^1\} n_{rl}^0 + \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < t_{(i)}^1\} \Bigg].
\end{aligned}
$$

$$
\begin{aligned}
I_2 =& P(T_{n_0+1}^0 < \alpha_1 L_x + \alpha_2 U_y) \frac{n_{ru}^1}{n_1+1} \\
\geq& \frac{n_{ru}^1}{n_1+1} \Bigg[\mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\} \frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 L_x + \alpha_2 U_y\} \frac{n_{lr}^0}{n_0+1} \\
&+ \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\} \frac{n_{rl}^0}{n_0+1} + \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < \alpha_1 L_x + \alpha_2 U_y\} \frac{1}{n_0+1} \\
&+ \mathbf{1}\{\infty < \alpha_1 L_x + \alpha_2 U_y\} \left(\frac{n_{lu}^0}{n_0+1} + \frac{n_{ru}^0}{n_0+1} + \frac{n_{ul}^0}{n_0+1} + \frac{n_{ur}^0}{n_0+1} + \frac{n_{uu}^0}{n_0+1} + \frac{1}{n_0+1} \right) \Bigg] \\
=& \frac{n_{ru}^1}{(n_1+1)(n_0+1)} \Bigg[n_{ll}^0 + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\} n_{rl}^0 + \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < \alpha_1 L_x + \alpha_2 U_y\} \Bigg].
\end{aligned}
$$

$$
\begin{aligned}
I_3 =& P(T_{n_0+1}^0 < \alpha_1 U_x + \alpha_2 L_y) \frac{n_{ur}^1}{n_1+1} \\
\geq& \frac{n_{ur}^1}{n_1+1} \Bigg[\mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y < \alpha_1 U_x + \alpha_2 L_y\} \frac{n_{ll}^0}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\} \frac{n_{lr}^0}{n_0+1} \\
&+ \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 U_x + \alpha_2 L_y\} \frac{n_{rl}^0}{n_0+1} + \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < \alpha_1 U_x + \alpha_2 L_y\} \frac{1}{n_0+1} \\
&+ \mathbf{1}\{\infty < \alpha_1 U_x + \alpha_2 L_y\} \left(\frac{n_{lu}^0}{n_0+1} + \frac{n_{ru}^0}{n_0+1} + \frac{n_{ul}^0}{n_0+1} + \frac{n_{ur}^0}{n_0+1} + \frac{n_{uu}^0}{n_0+1} + \frac{1}{n_0+1} \right) \Bigg] \\
=& \frac{n_{ur}^1}{(n_1+1)(n_0+1)} \Bigg[n_{ll}^0 + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\} n_{lr}^0 + \sum_{j=0}^{r_T^0-1} \mathbf{1}\{t_{(j+1)}^0 < \alpha_1 U_x + \alpha_2 L_y\} \Bigg].
\end{aligned}
$$

$$
\begin{aligned}
I_4 =& P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 U_y)\frac{n^1_{uu}}{n_1+1} \\
\geq& \frac{n^1_{uu}}{n_1+1}\Bigg[\mathbf{1}\{\alpha_1 L_x + \alpha_2 L_y < \alpha_1 U_x + \alpha_2 U_y\}\frac{n^0_{ll}}{n_0+1} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 U_y\}\frac{n^0_{lr}}{n_0+1} \\
&+\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 U_x + \alpha_2 U_y\}\frac{n^0_{rl}}{n_0+1} + \sum_{j=0}^{r^0_T-1}\mathbf{1}\{t^0_{(j+1)} < \alpha_1 U_x + \alpha_2 U_y\}\frac{1}{n_0+1} \\
&+\mathbf{1}\{\infty < \alpha_1 U_x + \alpha_2 U_y\}\left(\frac{n^0_{lu}}{n_0+1} + \frac{n^0_{ru}}{n_0+1} + \frac{n^0_{ul}}{n_0+1} + \frac{n^0_{ur}}{n_0+1} + \frac{n^0_{uu}}{n_0+1} + \frac{1}{n_0+1}\right)\Bigg] \\
=& \frac{n^1_{uu}}{n_1+1}\left[\frac{n^0_{ll}}{n_0+1} + \frac{n^0_{lr}}{n_0+1} + \frac{n^0_{rl}}{n_0+1} + \sum_{j=0}^{r^0_T-1}\mathbf{1}\{t^0_{(j+1)} < \alpha_1 U_x + \alpha_2 U_y\}\frac{1}{n_0+1}\right] \\
=& \frac{n^1_{uu}}{(n_1+1)(n_0+1)}\left[n^0_{ll} + n^0_{lr} + n^0_{rl} + n^0_{rr}\right].
\end{aligned}
$$

then the lower probability can be written as ($\underline{P} = I_1 + I_2 + I_3 + I_4$)

$$
\begin{aligned}
\underline{P} =& \frac{1}{(n_1+1)(n_0+1)}\sum_{i=1}^{r^1_T}\left[n^0_{ll} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t^1_{(i)}\}n^0_{lr} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t^1_{(i)}\}n^0_{rl} + \sum_{j=0}^{r^0_T-1}\mathbf{1}\{t^0_{(j+1)} < t^1_{(i)}\}\right] \\
&+\frac{n^1_{ru}}{(n_1+1)(n_0+1)}\left[n^0_{ll} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}n^0_{rl} + \sum_{j=0}^{r^0_T-1}\mathbf{1}\{t^0_{(j+1)} < \alpha_1 L_x + \alpha_2 U_y\}\right] \\
&+\frac{n^1_{ur}}{(n_1+1)(n_0+1)}\left[n^0_{ll} + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}n^0_{lr} + \sum_{j=0}^{r^0_T-1}\mathbf{1}\{t^0_{(j+1)} < \alpha_1 U_x + \alpha_2 L_y\}\right] \\
&+\frac{n^1_{uu}}{(n_1+1)(n_0+1)}\left[n^0_{ll} + n^0_{lr} + n^0_{rl} + n^0_{rr}\right].
\end{aligned}
$$

And the NPI upper probability for the event $T^0_{n_0+1} < T^1_{n_1+1}$ is obtained as follows:

$$
\begin{aligned}
P \leq& P(T^0_{n_0+1} < \infty)\left[\frac{n^1_{lu}}{n_1+1} + \frac{n^1_{ru}}{n_1+1} + \frac{n^1_{ul}}{n_1+1} + \frac{n^1_{ur}}{n_1+1} + \frac{n^1_{uu}}{n_1+1} + \frac{1}{n_1+1}\right] + \sum_{i=1}^{r^1_T}P(T^0_{n_0+1} < t^1_{(i)})\frac{1}{n_1+1} \\
&+ P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 L_y)\frac{n^1_{ll}}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 U_y)\frac{n^1_{lr}}{n_1+1} + P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 L_y)\frac{n^1_{rl}}{n_1+1}.
\end{aligned}
$$

The above inequality follows by putting all probability masses for $T^1_{n_1+1}$, according to the M-functions in Definition 1, corresponding to the intervals $(t^1_{(i)}, t^1_{(i+1)})$ $(i = 0, ..., r^1_T)$, $(-\infty, \alpha_1 L_x + \alpha_2 L_y)$, $(-\infty, \alpha_1 L_x + \alpha_2 U_y)$, $(-\infty, \alpha_1 U_x + \alpha_2 L_y)$,$(\alpha_1 L_x + \alpha_2 U_y, \infty)$, $(\alpha_1 U_x + \alpha_2 L_y, \infty)$, $(\alpha_1 U_x + \alpha_2 U_y, \infty)$ and $(-\infty, \infty)$ to the right-end points of these intervals. The next step in the proof is divided into five parts, J_1, J_2, J_3, J_4 and J_5. The following inequalities (in I_1, I_2, I_3 and I_4) follow by putting all probability masses for $T^0_{n_0+1}$ corresponding to

the intervals $(t^0_{(j)}, t^0_{(j+1)})$ $(j = 0, ..., r^0_T)$, $(-\infty, \alpha_1 L_x + \alpha_2 L_y)$, $(-\infty, \alpha_1 L_x + \alpha_2 U_y)$, $(-\infty, \alpha_1 U_x + \alpha_2 L_y)$,$(\alpha_1 L_x + \alpha_2 U_y, \infty)$, $(\alpha_1 U_x + \alpha_2 L_y, \infty)$, $(\alpha_1 U_x + \alpha_2 U_y, \infty)$ and $(-\infty, \infty)$ to the left-end points of these intervals.

$$
\begin{aligned}
J_1 =& P(T^0_{n_0+1} < \infty)\left[\frac{n^1_{lu}}{n_1+1} + \frac{n^1_{ru}}{n_1+1} + \frac{n^1_{ul}}{n_1+1} + \frac{n^1_{ur}}{n_1+1} + \frac{n^1_{uu}}{n_1+1} + \frac{1}{n_1+1}\right] \\
=& \frac{n^1_{lu}}{n_1+1} + \frac{n^1_{ru}}{n_1+1} + \frac{n^1_{ul}}{n_1+1} + \frac{n^1_{ur}}{n_1+1} + \frac{n^1_{uu}}{n_1+1} + \frac{1}{n_1+1}. \\
J_2 =& \sum_{i=1}^{r^1_T} P(T^0_{n_0+1} < t^1_{(i)}) \frac{1}{n_1+1} \\
=& \sum_{i=1}^{r^1_T} \frac{1}{n_1+1}\left[\mathbf{1}\{-\infty < t^1_{(i)}\}\left[\frac{n^0_{ll}}{n_0+1} + \frac{n^0_{lr}}{n_0+1} + \frac{n^0_{lu}}{n_0+1} + \frac{n^0_{rl}}{n_0+1} + \frac{n^0_{ul}}{n_0+1}\right] + \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t^1_{(i)}\}\frac{n^0_{ru}}{n_0+1}\right. \\
&\left. + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t^1_{(i)}\}\frac{n^0_{ur}}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y < t^1_{(i)}\}\frac{n^0_{uu}}{n_0+1} + \sum_{j=0}^{r^0_T} \mathbf{1}\{t^0_{(j)} < t^1_{(i)}\}\frac{1}{n_0+1}\right] \\
=& \frac{n^1_{rr}}{n_1+1}\left[\frac{l^0_x + l^0_y - n^0_{ll} + 1}{n_0+1}\right] + \frac{1}{(n_1+1)(n_0+1)} \sum_{i=1}^{r^1_T}\left[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t^1_{(i)}\} n^0_{ru} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t^1_{(i)}\} n^0_{ur}\right. \\
&\left. + \sum_{j=1}^{r^0_T} \mathbf{1}\{t^0_{(j)} < t^1_{(i)}\}\right].
\end{aligned}
$$

$$
\begin{aligned}
J_3 &= P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 L_y)\frac{n^1_{ll}}{n_1+1} \\
&= \frac{n^1_{ll}}{n_1+1}\left[\mathbf{1}\{-\infty < \alpha_1 L_x + \alpha_2 L_y\}\left\{\frac{n^0_{ll}}{n_0+1} + \frac{n^0_{lr}}{n_0+1} + \frac{n^0_{lu}}{n_0+1} + \frac{n^0_{rl}}{n_0+1} + \frac{n^0_{ul}}{n_0+1}\right\}\right. \\
&+ \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 L_x + \alpha_2 L_y\}\frac{n^0_{ru}}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 L_y\}\frac{n^0_{ur}}{n_0+1} \\
&\left. + \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y < \alpha_1 L_x + \alpha_2 L_y\}\frac{n^0_{uu}}{n_0+1} + \sum_{j=0}^{r^0_T} \mathbf{1}\{t^0_{(j)} < \alpha_1 L_x + \alpha_2 L_y\}\frac{1}{n_0+1}\right] \\
&= \frac{n^1_{ll}}{n_1+1}\left\{\frac{l^0_x + l^0_y - n^0_{ll} + 1}{n_0+1}\right\}.
\end{aligned}
$$

$$
\begin{aligned}
J_4 &= P(T^0_{n_0+1} < \alpha_1 L_x + \alpha_2 U_y)\frac{n^1_{lr}}{n_1+1} \\
&= \frac{n^1_{lr}}{n_1+1}\Bigg[\mathbf{1}\{-\infty < \alpha_1 L_x + \alpha_2 U_y\}\left\{\frac{n^0_{ll}}{n_0+1} + \frac{n^0_{lr}}{n_0+1} + \frac{n^0_{lu}}{n_0+1} + \frac{n^0_{rl}}{n_0+1} + \frac{n^0_{ul}}{n_0+1}\right\} \\
&+ \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n^0_{ru}}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n^0_{ur}}{n_0+1} \\
&+ \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n^0_{uu}}{n_0+1} + \sum_{j=0}^{r^0_T}\mathbf{1}\{t^0_{(j)} < \alpha_1 L_x + \alpha_2 U_y\}\frac{1}{n_0+1}\Bigg] \\
&= \frac{n^1_{lr}}{n_1+1}\left\{\frac{l^0_x + l^0_y - n^0_{ll} + 1}{n_0+1}\right\} + \frac{n^1_{lr}}{n_0+1}\Bigg[\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n^0_{ur}}{n_0+1} \\
&+ \sum_{j=1}^{r^0_T}\mathbf{1}\{t^0_{(j)} < \alpha_1 L_x + \alpha_2 U_y\}\frac{1}{n_0+1}\Bigg].
\end{aligned}
$$

$$
\begin{aligned}
J_5 &= P(T^0_{n_0+1} < \alpha_1 U_x + \alpha_2 L_y)\frac{n^1_{rl}}{n_1+1} \\
&= \frac{n^1_{rl}}{n_1+1}\Bigg[\mathbf{1}\{-\infty < \alpha_1 U_x + \alpha_2 L_y\}\left\{\frac{n^0_{ll}}{n_0+1} + \frac{n^0_{lr}}{n_0+1} + \frac{n^0_{lu}}{n_0+1} + \frac{n^0_{rl}}{n_0+1} + \frac{n^0_{ul}}{n_0+1}\right\} \\
&+ \mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n^0_{ru}}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n^0_{ur}}{n_0+1} \\
&+ \mathbf{1}\{\alpha_1 U_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n^0_{uu}}{n_0+1} + \sum_{j=0}^{r^0_T}\mathbf{1}\{t^0_{(j)} < \alpha_1 U_x + \alpha_2 L_y\}\frac{1}{n_0+1}\Bigg] \\
&= \frac{n^1_{rl}}{n_1+1}\left\{\frac{l^0_x + l^0_y - n^0_{ll} + 1}{n_0+1}\right\} + \frac{n^1_{rl}}{n_1+1}\Bigg[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n^0_{ru}}{n_0+1} \\
&+ \sum_{j=1}^{r^0_T}\mathbf{1}\{t^0_{(j)} < \alpha_1 U_x + \alpha_2 L_y\}\frac{1}{n_0+1}\Bigg].
\end{aligned}
$$

then the upper probability can be written as ($\overline{P} = J_1 + J_2 + J_3 + J_4 + J_5$)

$$\overline{P} = \frac{u_x^1 + u_y^1 - n_{uu}^1 + 1}{n_1 + 1} + \left[\frac{n_{rr}^1}{n_1+1} + \frac{n_{ll}^1}{n_1+1} + \frac{n_{lr}^1}{n_1+1} + \frac{n_{rl}^1}{n_1+1}\right]\left\{\frac{l_x^0 + l_y^0 - n_{ll}^0 + 1}{n_0+1}\right\}$$
$$+ \sum_{i=1}^{r_T^1} \frac{1}{n_1+1}\left[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < t_{(i)}^1\}\frac{n_{ru}^0}{n_0+1} + \mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < t_{(i)}^1\}\frac{n_{ur}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < t_{(i)}^1\}\frac{1}{n_0+1}\right]$$
$$+ \frac{n_{lr}^1}{n_1+1}\left[\mathbf{1}\{\alpha_1 U_x + \alpha_2 L_y < \alpha_1 L_x + \alpha_2 U_y\}\frac{n_{ur}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < \alpha_1 L_x + \alpha_2 U_y\}\frac{1}{n_0+1}\right]$$
$$+ \frac{n_{rl}^1}{n_1+1}\left[\mathbf{1}\{\alpha_1 L_x + \alpha_2 U_y < \alpha_1 U_x + \alpha_2 L_y\}\frac{n_{ru}^0}{n_0+1} + \sum_{j=1}^{r_T^0}\mathbf{1}\{t_{(j)}^0 < \alpha_1 U_x + \alpha_2 L_y\}\frac{1}{n_0+1}\right].$$

A2. Data structure for combining three and four biomarkers

If one wants to combine three biomarkers, say $T_{n+1} = \alpha_1 Z_{1,n+1} + \alpha_2 Z_{2,n+1} + \alpha_3 Z_{3,n+1}$, then the two matrices Q_L and Q_U will be of order $(r_T^0 + 27) \times (r_T^1 + 27)$. The vector M^D, with the M-functions as in Definition 1, can be defined using the following data structure,

$$\mathbb{S} = \begin{array}{c} \\ q_{uu.} \\ q_{ur.} \\ q_{ul.} \\ q_{ru.} \\ q_{rr.} \\ q_{rl.} \\ q_{lu.} \\ q_{lr.} \\ q_{ll.} \end{array} \begin{array}{c} \begin{array}{ccc} q_{..l} & q_{..r} & q_{..u} \end{array} \\ \left(\begin{array}{ccc} n_{uul} & n_{uur} & n_{uuu} \\ n_{url} & n_{urr} & n_{uru} \\ n_{ull} & n_{ulr} & n_{ulu} \\ \hline n_{rul} & n_{rur} & n_{ruu} \\ n_{rrl} & n_{rrr} & n_{rru} \\ n_{rll} & n_{rlr} & n_{rlu} \\ \hline n_{lul} & n_{lur} & n_{luu} \\ n_{lrl} & n_{lrr} & n_{lru} \\ n_{lll} & n_{llr} & n_{llu} \end{array}\right) \end{array},$$

where

$$l_{z_1} = q_{ll.} + q_{lr.} + q_{lu.}, \qquad l_{z_2} = q_{ll.} + q_{rl.} + q_{ul.}, \qquad l_{z_3} = q_{..l},$$
$$r_{z_1} = q_{rl.} + q_{rr.} + q_{ru.}, \qquad r_{z_2} = q_{lr.} + q_{rr.} + q_{ur.}, \qquad r_{z_3} = q_{..r},$$
$$u_{z_1} = q_{ul.} + q_{ur.} + q_{uu.}, \quad u_{z_2} = q_{lu.} + q_{ru.} + q_{uu.}, \quad u_{z_3} = q_{..u}.$$

If one wants to combine four biomarkers, say $T_{n+1} = \alpha_1 Z_{1,n+1} + \alpha_2 Z_{2,n+1} + \alpha_3 Z_{3,n+1} + \alpha_4 Z_{4,n+1}$, then the two matrices Q_L and Q_U will be of order $(r_T^0 + 81) \times (r_T^1 + 81)$. The vector M^D, with M-functions as in

Definition 1, can be defined using the following data structure,

$$\mathbb{S} = \begin{array}{c} \\ q_{.uu.} \\ q_{.ur.} \\ q_{.ul.} \\ q_{.ru.} \\ q_{.rr.} \\ q_{.rl.} \\ q_{.lu.} \\ q_{.lr.} \\ q_{.ll.} \end{array} \begin{array}{c} \begin{array}{ccc|ccc|ccc} q_{l..l} & q_{l..r} & q_{l..u} & q_{r..l} & q_{r..r} & q_{r..u} & q_{u..l} & q_{u..r} & q_{u..u} \end{array} \\ \left(\begin{array}{ccc|ccc|ccc} n_{luul} & n_{luur} & n_{luuu} & n_{ruul} & n_{ruur} & n_{ruuu} & n_{uuul} & n_{uuur} & n_{uuuu} \\ n_{lurl} & n_{lurr} & n_{luru} & n_{rurl} & n_{rurr} & n_{ruru} & n_{uurl} & n_{uurr} & n_{uuru} \\ n_{lull} & n_{lulr} & n_{lulu} & n_{rull} & n_{rulr} & n_{rulu} & n_{uull} & n_{uulr} & n_{uulu} \\ \hline n_{lrul} & n_{lrur} & n_{lruu} & n_{rrul} & n_{rrur} & n_{rruu} & n_{urul} & n_{urur} & n_{uruu} \\ n_{lrrl} & n_{lrrr} & n_{lrru} & n_{rrrl} & n_{rrrr} & n_{rrru} & n_{urrl} & n_{urrr} & n_{urru} \\ n_{lrll} & n_{lrlr} & n_{lrlu} & n_{rrll} & n_{rrlr} & n_{rrlu} & n_{urll} & n_{urlr} & n_{urlu} \\ \hline n_{llul} & n_{llur} & n_{lluu} & n_{rlul} & n_{rlur} & n_{rluu} & n_{ulul} & n_{ulur} & n_{uluu} \\ n_{llrl} & n_{llrr} & n_{llru} & n_{rlrl} & n_{rlrr} & n_{rlru} & n_{ulrl} & n_{ulrr} & n_{ulru} \\ n_{llll} & n_{lllr} & n_{lllu} & n_{rlll} & n_{rllr} & n_{rllu} & n_{ulll} & n_{ullr} & n_{ullu} \end{array}\right) \end{array},$$

where

$$l_{z_1} = q_{l..l} + q_{l..r} + q_{l..u}, \quad l_{z_2} = q_{.ll.} + q_{.lr.} + q_{.lu.}, \quad l_{z_3} = q_{.ll.} + q_{.rl.} + q_{.ul.}, \quad l_{z_4} = q_{l..l} + q_{r..l} + q_{u..l},$$
$$r_{z_1} = q_{r..l} + q_{r..r} + q_{r..u}, \quad r_{z_2} = q_{.rl.} + q_{.rr.} + q_{.ru.}, \quad r_{z_3} = q_{.lr.} + q_{.rr.} + q_{.ur.}, \quad r_{z_4} = q_{l..r} + q_{r..r} + q_{u..r},$$
$$u_{z_1} = q_{u..l} + q_{u..r} + q_{u..u}, \quad u_{z_2} = q_{.ul.} + q_{.ur.} + q_{.uu.}, \quad u_{z_3} = q_{.lu.} + q_{.ru.} + q_{.uu.}, \quad u_{z_4} = q_{l..u} + q_{r..u} + q_{u..u}.$$

A3. Simulation results, Section 6.2

μ_x	μ_y	AUC_x	AUC_y	$\hat{\alpha}_{opt}$		$\widehat{AUC}$	α^L_{opt}		$\underline{AUC}$	α^U_{opt}		$\overline{AUC}$
$n_0 = n_1 = 50$, $\rho = 0$												
0.358	0.358	0.6	0.6	0.505	0.495	0.655	0.506	0.494	0.630	0.506	0.494	0.669
0.742	0.358	0.7	0.6	0.680	0.320	0.729	0.680	0.320	0.701	0.681	0.319	0.740
0.742	0.742	0.7	0.7	0.499	0.501	0.779	0.500	0.500	0.748	0.500	0.500	0.787
1.190	0.358	0.8	0.6	0.775	0.225	0.816	0.777	0.223	0.785	0.775	0.225	0.823
1.190	0.742	0.8	0.7	0.622	0.378	0.845	0.623	0.377	0.812	0.624	0.376	0.851
1.190	1.190	0.8	0.8	0.500	0.500	0.887	0.501	0.499	0.853	0.500	0.500	0.892
1.812	0.358	0.9	0.6	0.837	0.163	0.908	0.837	0.163	0.872	0.837	0.163	0.911
1.812	0.742	0.9	0.7	0.719	0.281	0.920	0.720	0.280	0.884	0.721	0.279	0.923
1.812	1.190	0.9	0.8	0.610	0.390	0.940	0.610	0.390	0.904	0.610	0.390	0.943
1.812	1.812	0.9	0.9	0.504	0.496	0.967	0.505	0.495	0.930	0.505	0.495	0.968
$n_0 = n_1 = 50$, $\rho = 0.5$												
0.358	0.358	0.6	0.6	0.500	0.500	0.630	0.502	0.498	0.606	0.502	0.498	0.645
0.742	0.358	0.7	0.6	0.866	0.340	0.705	0.866	0.134	0.678	0.867	0.133	0.716
0.742	0.742	0.7	0.7	0.504	0.496	0.737	0.504	0.496	0.708	0.505	0.495	0.747
1.190	0.358	0.8	0.6	0.974	0.026	0.800	0.974	0.026	0.769	0.974	0.026	0.808
1.190	0.742	0.8	0.7	0.818	0.182	0.808	0.820	0.180	0.776	0.820	0.180	0.815
1.190	1.190	0.8	0.8	0.504	0.496	0.840	0.507	0.493	0.807	0.506	0.494	0.846
1.812	0.358	0.9	0.6	0.994	0.006	0.899	0.995	0.005	0.864	0.995	0.005	0.903
1.812	0.742	0.9	0.7	0.959	0.041	0.900	0.961	0.039	0.865	0.960	0.040	0.904
1.812	1.190	0.9	0.8	0.805	0.195	0.907	0.806	0.194	0.872	0.807	0.193	0.910
1.812	1.812	0.9	0.9	0.504	0.496	0.933	0.507	0.493	0.897	0.507	0.493	0.936
$n_0 = n_1 = 50$, $\rho = 0.75$												
0.358	0.358	0.6	0.6	0.505	0.495	0.620	0.506	0.494	0.596	0.506	0.494	0.634
0.742	0.358	0.7	0.6	0.960	0.040	0.700	0.960	0.040	0.673	0.961	0.039	0.711
0.742	0.742	0.7	0.7	0.502	0.498	0.721	0.505	0.495	0.693	0.503	0.497	0.732
1.190	0.358	0.8	0.6	0.997	0.003	0.799	0.998	0.002	0.768	0.998	0.002	0.807
1.190	0.742	0.8	0.7	0.952	0.048	0.800	0.954	0.046	0.769	0.954	0.046	0.808
1.190	1.190	0.8	0.8	0.503	0.497	0.822	0.503	0.497	0.790	0.504	0.496	0.828
1.812	0.358	0.9	0.6	1.000	0.000	0.899	1.000	0.000	0.864	1.000	0.000	0.903
1.812	0.742	0.9	0.7	0.997	0.003	0.899	0.998	0.002	0.864	0.998	0.002	0.903
1.812	1.190	0.9	0.8	0.960	0.040	0.900	0.963	0.037	0.865	0.963	0.037	0.904
1.812	1.812	0.9	0.9	0.509	0.491	0.918	0.515	0.485	0.882	0.514	0.486	0.921
$n_0 = n_1 = 100$, $\rho = 0$												
0.358	0.358	0.6	0.6	0.494	0.506	0.647	0.494	0.506	0.634	0.494	0.506	0.654
0.742	0.358	0.7	0.6	0.678	0.322	0.724	0.678	0.322	0.709	0.678	0.322	0.729
0.742	0.742	0.7	0.7	0.498	0.502	0.774	0.498	0.502	0.759	0.497	0.503	0.779
1.190	0.358	0.8	0.6	0.772	0.228	0.812	0.773	0.227	0.796	0.772	0.228	0.816
1.190	0.742	0.8	0.7	0.617	0.383	0.841	0.617	0.383	0.825	0.617	0.383	0.845
1.190	1.190	0.8	0.8	0.498	0.502	0.885	0.498	0.502	0.868	0.498	0.502	0.887
1.812	0.358	0.9	0.6	0.839	0.161	0.905	0.840	0.160	0.888	0.840	0.160	0.907
1.812	0.742	0.9	0.7	0.711	0.289	0.918	0.711	0.289	0.900	0.711	0.289	0.920
1.812	1.190	0.9	0.8	0.604	0.396	0.939	0.603	0.397	0.920	0.604	0.396	0.940
1.812	1.812	0.9	0.9	0.499	0.501	0.966	0.499	0.501	0.947	0.499	0.501	0.967
$n_0 = n_1 = 100$, $\rho = 0.5$												
0.358	0.358	0.6	0.6	0.510	0.490	0.622	0.511	0.489	0.609	0.511	0.489	0.629
0.742	0.358	0.7	0.6	0.916	0.084	0.702	0.916	0.084	0.688	0.916	0.084	0.707
0.742	0.742	0.7	0.7	0.508	0.492	0.731	0.508	0.492	0.716	0.508	0.492	0.736
1.190	0.358	0.8	0.6	0.991	0.009	0.799	0.991	0.009	0.784	0.992	0.008	0.803
1.190	0.742	0.8	0.7	0.841	0.159	0.804	0.841	0.159	0.789	0.840	0.160	0.808
1.190	1.190	0.8	0.8	0.504	0.496	0.836	0.504	0.496	0.820	0.505	0.495	0.839
1.812	0.358	0.9	0.6	0.999	0.001	0.899	0.999	0.001	0.882	0.999	0.001	0.901
1.812	0.742	0.9	0.7	0.980	0.020	0.900	0.980	0.020	0.882	0.980	0.020	0.902
1.812	1.190	0.9	0.8	0.817	0.183	0.905	0.817	0.183	0.887	0.817	0.183	0.907
1.812	1.812	0.9	0.9	0.504	0.496	0.931	0.505	0.495	0.913	0.504	0.496	0.933
$n_0 = n_1 = 100$, $\rho = 0.75$												
0.358	0.358	0.6	0.6	0.492	0.508	0.613	0.492	0.508	0.600	0.491	0.509	0.620
0.742	0.358	0.7	0.6	0.989	0.011	0.698	0.989	0.011	0.684	0.989	0.011	0.704
0.742	0.742	0.7	0.7	0.493	0.507	0.716	0.494	0.506	0.702	0.494	0.506	0.722
1.190	0.358	0.8	0.6	1.000	0.000	0.798	1.000	0.000	0.782	1.000	0.000	0.802
1.190	0.742	0.8	0.7	0.981	0.019	0.798	0.981	0.019	0.783	0.981	0.019	0.802
1.190	1.190	0.8	0.8	0.492	0.508	0.818	0.493	0.507	0.802	0.492	0.508	0.821
1.812	0.358	0.9	0.6	1.000	0.000	0.899	1.000	0.000	0.881	1.000	0.000	0.901
1.812	0.742	0.9	0.7	0.999	0.001	0.899	0.999	0.001	0.881	0.999	0.001	0.901
1.812	1.190	0.9	0.8	0.981	0.019	0.899	0.981	0.019	0.881	0.980	0.020	0.901
1.812	1.812	0.9	0.9	0.496	0.504	0.916	0.496	0.504	0.898	0.497	0.503	0.917

Table 5: Simulated data example, Bivariate normal distribution, without LoD

μ_x	μ_y	AUC_x	AUC_y	$\hat{\alpha}_{opt}$		$\widehat{AUC}$	α^L_{opt}		$\underline{AUC}$	α^U_{opt}		$\overline{AUC}$
$n_0 = n_1 = 50$, $\rho = 0$												
0.358	0.358	0.6	0.6	0.510	0.490	0.628	0.506	0.494	0.439	0.506	0.494	0.783
0.742	0.358	0.7	0.6	0.667	0.333	0.681	0.613	0.387	0.499	0.678	0.322	0.829
0.742	0.742	0.7	0.7	0.505	0.495	0.719	0.502	0.498	0.550	0.503	0.497	0.857
1.190	0.358	0.8	0.6	0.773	0.227	0.759	0.623	0.377	0.570	0.776	0.224	0.884
1.190	0.742	0.8	0.7	0.635	0.365	0.783	0.548	0.452	0.616	0.641	0.359	0.897
1.190	1.190	0.8	0.8	0.504	0.496	0.826	0.501	0.499	0.673	0.501	0.499	0.921
1.812	0.358	0.9	0.6	0.825	0.175	0.865	0.575	0.425	0.646	0.835	0.165	0.942
1.812	0.742	0.9	0.7	0.729	0.271	0.876	0.530	0.470	0.691	0.738	0.262	0.947
1.812	1.190	0.9	0.8	0.620	0.380	0.897	0.510	0.490	0.741	0.624	0.376	0.957
1.812	1.812	0.9	0.9	0.503	0.497	0.937	0.502	0.498	0.797	0.506	0.494	0.974
$n_0 = n_1 = 50$, $\rho = 0.5$												
0.358	0.358	0.6	0.6	0.518	0.482	0.604	0.497	0.503	0.465	0.509	0.491	0.742
0.742	0.358	0.7	0.6	0.807	0.193	0.658	0.739	0.261	0.525	0.821	0.179	0.795
0.742	0.742	0.7	0.7	0.505	0.495	0.672	0.498	0.502	0.563	0.501	0.499	0.812
1.190	0.358	0.8	0.6	0.936	0.064	0.748	0.799	0.201	0.600	0.935	0.065	0.867
1.190	0.742	0.8	0.7	0.802	0.198	0.740	0.633	0.367	0.628	0.801	0.199	0.863
1.190	1.190	0.8	0.8	0.514	0.486	0.763	0.503	0.497	0.675	0.508	0.492	0.879
1.812	0.358	0.9	0.6	0.976	0.024	0.868	0.759	0.241	0.674	0.974	0.026	0.938
1.812	0.742	0.9	0.7	0.935	0.065	0.856	0.640	0.360	0.704	0.933	0.067	0.932
1.812	1.190	0.9	0.8	0.801	0.199	0.851	0.557	0.443	0.744	0.799	0.201	0.929
1.812	1.812	0.9	0.9	0.506	0.494	0.881	0.501	0.499	0.797	0.506	0.494	0.944
$n_0 = n_1 = 50$, $\rho = 0.75$												
0.358	0.358	0.6	0.6	0.503	0.497	0.593	0.514	0.486	0.493	0.507	0.493	0.712
0.742	0.358	0.7	0.6	0.926	0.074	0.661	0.857	0.143	0.562	0.923	0.077	0.778
0.742	0.742	0.7	0.7	0.508	0.492	0.660	0.506	0.494	0.590	0.503	0.497	0.785
1.190	0.358	0.8	0.6	0.992	0.008	0.770	0.927	0.073	0.643	0.990	0.010	0.865
1.190	0.742	0.8	0.7	0.931	0.069	0.742	0.749	0.251	0.658	0.919	0.081	0.850
1.190	1.190	0.8	0.8	0.504	0.496	0.746	0.505	0.495	0.696	0.496	0.504	0.856
1.812	0.358	0.9	0.6	0.999	0.001	0.894	0.917	0.083	0.717	0.998	0.002	0.945
1.812	0.742	0.9	0.7	0.993	0.007	0.873	0.799	0.201	0.737	0.992	0.008	0.934
1.812	1.190	0.9	0.8	0.948	0.052	0.851	0.628	0.372	0.765	0.940	0.060	0.922
1.812	1.812	0.9	0.9	0.510	0.490	0.862	0.504	0.496	0.812	0.514	0.486	0.929
$n_0 = n_1 = 100$, $\rho = 0$												
0.358	0.358	0.6	0.6	0.490	0.510	0.613	0.498	0.502	0.439	0.497	0.503	0.771
0.742	0.358	0.7	0.6	0.681	0.319	0.669	0.611	0.389	0.503	0.684	0.316	0.821
0.742	0.742	0.7	0.7	0.499	0.501	0.708	0.503	0.497	0.556	0.501	0.499	0.849
1.190	0.358	0.8	0.6	0.774	0.226	0.752	0.608	0.392	0.557	0.776	0.224	0.879
1.190	0.742	0.8	0.7	0.625	0.375	0.774	0.542	0.458	0.625	0.628	0.372	0.892
1.190	1.190	0.8	0.8	0.506	0.494	0.818	0.501	0.499	0.684	0.501	0.499	0.917
1.812	0.358	0.9	0.6	0.834	0.166	0.861	0.565	0.435	0.655	0.838	0.162	0.940
1.812	0.742	0.9	0.7	0.723	0.277	0.871	0.524	0.476	0.702	0.726	0.274	0.944
1.812	1.190	0.9	0.8	0.613	0.387	0.893	0.504	0.496	0.755	0.615	0.385	0.954
1.812	1.812	0.9	0.9	0.502	0.498	0.934	0.502	0.498	0.812	0.502	0.498	0.972
$n_0 = n_1 = 100$, $\rho = 0.5$												
0.358	0.358	0.6	0.6	0.505	0.495	0.589	0.503	0.497	0.465	0.505	0.495	0.727
0.742	0.358	0.7	0.6	0.861	0.139	0.648	0.756	0.244	0.532	0.862	0.138	0.787
0.742	0.742	0.7	0.7	0.494	0.506	0.663	0.499	0.501	0.569	0.505	0.495	0.802
1.190	0.358	0.8	0.6	0.969	0.031	0.745	0.805	0.195	0.609	0.965	0.035	0.862
1.190	0.742	0.8	0.7	0.822	0.178	0.734	0.621	0.379	0.637	0.810	0.190	0.857
1.190	1.190	0.8	0.8	0.499	0.501	0.757	0.501	0.499	0.686	0.502	0.498	0.872
1.812	0.358	0.9	0.6	0.993	0.007	0.867	0.761	0.239	0.685	0.992	0.008	0.937
1.812	0.742	0.9	0.7	0.962	0.038	0.855	0.626	0.374	0.716	0.959	0.041	0.930
1.812	1.190	0.9	0.8	0.809	0.191	0.848	0.551	0.449	0.757	0.802	0.198	0.926
1.812	1.812	0.9	0.9	0.503	0.497	0.877	0.500	0.500	0.812	0.503	0.497	0.941
$n_0 = n_1 = 100$, $\rho = 0.75$												
0.358	0.358	0.6	0.6	0.486	0.514	0.580	0.503	0.497	0.496	0.496	0.504	0.696
0.742	0.358	0.7	0.6	0.973	0.027	0.656	0.894	0.106	0.570	0.965	0.035	0.770
0.742	0.742	0.7	0.7	0.494	0.506	0.650	0.498	0.502	0.596	0.490	0.510	0.774
1.190	0.358	0.8	0.6	0.998	0.002	0.768	0.958	0.042	0.656	0.997	0.003	0.860
1.190	0.742	0.8	0.7	0.965	0.035	0.738	0.746	0.254	0.668	0.951	0.049	0.843
1.190	1.190	0.8	0.8	0.503	0.497	0.740	0.498	0.502	0.707	0.496	0.504	0.849
1.812	0.358	0.9	0.6	1.000	0.000	0.895	0.952	0.048	0.732	1.000	0.000	0.944
1.812	0.742	0.9	0.7	0.998	0.002	0.872	0.805	0.195	0.751	0.998	0.002	0.932
1.812	1.190	0.9	0.8	0.972	0.028	0.849	0.612	0.388	0.778	0.965	0.035	0.919
1.812	1.812	0.9	0.9	0.498	0.502	0.858	0.499	0.501	0.827	0.500	0.500	0.925

Table 6: Simulated data example, Bivariate normal distribution, with LoD 10%

β_x^0	β_y^0	AUC_x	AUC_y	$\hat{\alpha}_{opt}$		$\widehat{AUC}$	α_{opt}^L		$\underline{AUC}$	α_{opt}^U		$\overline{AUC}$
$n_0 = n_1 = 50$, $\rho = 0$												
0.667	0.667	0.6	0.6	0.516	0.484	0.663	0.518	0.482	0.638	0.518	0.482	0.676
0.429	0.667	0.7	0.6	0.725	0.275	0.739	0.726	0.274	0.710	0.726	0.274	0.749
0.429	0.429	0.7	0.7	0.509	0.491	0.791	0.510	0.490	0.761	0.510	0.490	0.799
0.250	0.667	0.8	0.6	0.850	0.150	0.824	0.851	0.149	0.792	0.851	0.149	0.831
0.250	0.429	0.8	0.7	0.692	0.308	0.857	0.694	0.306	0.824	0.694	0.306	0.863
0.250	0.250	0.8	0.8	0.507	0.493	0.900	0.510	0.490	0.865	0.510	0.490	0.904
0.111	0.667	0.9	0.6	0.936	0.064	0.912	0.937	0.063	0.877	0.937	0.063	0.916
0.111	0.429	0.9	0.7	0.856	0.144	0.928	0.858	0.142	0.892	0.858	0.142	0.931
0.111	0.250	0.9	0.8	0.731	0.269	0.949	0.733	0.267	0.912	0.733	0.267	0.951
0.111	0.111	0.9	0.9	0.511	0.489	0.973	0.514	0.486	0.936	0.514	0.486	0.974
$n_0 = n_1 = 50$, $\rho = 0.5$												
0.667	0.667	0.6	0.6	0.514	0.486	0.635	0.516	0.484	0.611	0.515	0.485	0.649
0.429	0.667	0.7	0.6	0.857	0.143	0.709	0.856	0.144	0.682	0.858	0.142	0.721
0.429	0.429	0.7	0.7	0.511	0.489	0.746	0.512	0.488	0.717	0.512	0.488	0.756
0.250	0.667	0.8	0.6	0.970	0.030	0.802	0.971	0.029	0.771	0.971	0.029	0.810
0.250	0.429	0.8	0.7	0.814	0.186	0.815	0.815	0.185	0.783	0.816	0.184	0.822
0.250	0.250	0.8	0.8	0.514	0.486	0.851	0.516	0.484	0.818	0.516	0.484	0.857
0.111	0.667	0.9	0.6	0.995	0.005	0.900	0.995	0.005	0.865	0.995	0.005	0.904
0.111	0.429	0.9	0.7	0.962	0.038	0.903	0.963	0.037	0.868	0.964	0.036	0.907
0.111	0.250	0.9	0.8	0.832	0.168	0.914	0.836	0.164	0.879	0.836	0.164	0.917
0.111	0.111	0.9	0.9	0.518	0.482	0.942	0.523	0.477	0.905	0.523	0.477	0.944
$n_0 = n_1 = 50$, $\rho = 0.75$												
0.667	0.667	0.6	0.6	0.517	0.483	0.622	0.517	0.483	0.598	0.518	0.482	0.637
0.429	0.667	0.7	0.6	0.955	0.045	0.702	0.956	0.044	0.674	0.956	0.044	0.713
0.429	0.429	0.7	0.7	0.507	0.493	0.726	0.509	0.491	0.698	0.509	0.491	0.737
0.250	0.667	0.8	0.6	0.997	0.003	0.800	0.997	0.003	0.769	0.997	0.003	0.807
0.250	0.429	0.8	0.7	0.935	0.065	0.802	0.937	0.063	0.771	0.937	0.063	0.810
0.250	0.250	0.8	0.8	0.519	0.481	0.828	0.522	0.478	0.796	0.521	0.479	0.835
0.111	0.667	0.9	0.6	1.0	0.0	0.9	1.000	0.000	0.865	1.000	0.000	0.903
0.111	0.429	0.9	0.7	0.996	0.004	0.900	0.997	0.003	0.865	0.997	0.003	0.904
0.111	0.250	0.9	0.8	0.944	0.056	0.902	0.946	0.054	0.867	0.947	0.053	0.906
0.111	0.111	0.9	0.9	0.514	0.486	0.923	0.520	0.480	0.888	0.521	0.479	0.926
$n_0 = n_1 = 100$, $\rho = 0$												
0.667	0.667	0.6	0.6	0.507	0.493	0.655	0.508	0.492	0.642	0.508	0.492	0.661
0.429	0.667	0.7	0.6	0.732	0.268	0.733	0.732	0.268	0.718	0.732	0.268	0.738
0.429	0.429	0.7	0.7	0.503	0.497	0.787	0.503	0.497	0.772	0.503	0.497	0.791
0.250	0.667	0.8	0.6	0.857	0.143	0.820	0.857	0.143	0.804	0.857	0.143	0.823
0.250	0.429	0.8	0.7	0.689	0.311	0.854	0.689	0.311	0.837	0.689	0.311	0.857
0.250	0.250	0.8	0.8	0.503	0.497	0.898	0.503	0.497	0.880	0.503	0.497	0.900
0.111	0.667	0.9	0.6	0.939	0.061	0.910	0.939	0.061	0.892	0.939	0.061	0.912
0.111	0.429	0.9	0.7	0.854	0.146	0.926	0.854	0.146	0.908	0.854	0.146	0.928
0.111	0.250	0.9	0.8	0.724	0.276	0.948	0.725	0.275	0.929	0.725	0.275	0.949
0.111	0.111	0.9	0.9	0.507	0.493	0.973	0.509	0.491	0.954	0.509	0.491	0.974
$n_0 = n_1 = 100$, $\rho = 0.5$												
0.667	0.667	0.6	0.6	0.508	0.492	0.628	0.508	0.492	0.615	0.508	0.492	0.635
0.429	0.667	0.7	0.6	0.891	0.109	0.705	0.892	0.108	0.691	0.891	0.109	0.711
0.429	0.429	0.7	0.7	0.507	0.493	0.741	0.507	0.493	0.727	0.507	0.493	0.746
0.250	0.667	0.8	0.6	0.984	0.016	0.801	0.984	0.016	0.785	0.984	0.016	0.805
0.250	0.429	0.8	0.7	0.826	0.174	0.812	0.826	0.174	0.796	0.826	0.174	0.815
0.250	0.250	0.8	0.8	0.505	0.495	0.849	0.505	0.495	0.832	0.504	0.496	0.852
0.111	0.667	0.9	0.6	0.998	0.002	0.900	0.998	0.002	0.882	0.998	0.002	0.902
0.111	0.429	0.9	0.7	0.971	0.029	0.902	0.972	0.028	0.884	0.972	0.028	0.904
0.111	0.250	0.9	0.8	0.832	0.168	0.913	0.833	0.167	0.895	0.833	0.167	0.914
0.111	0.111	0.9	0.9	0.511	0.489	0.942	0.512	0.488	0.923	0.512	0.488	0.943
$n_0 = n_1 = 100$, $\rho = 0.75$												
0.667	0.667	0.6	0.6	0.507	0.493	0.616	0.507	0.493	0.604	0.507	0.493	0.624
0.429	0.667	0.7	0.6	0.983	0.017	0.700	0.984	0.016	0.686	0.984	0.016	0.706
0.429	0.429	0.7	0.7	0.505	0.495	0.722	0.505	0.495	0.708	0.505	0.495	0.727
0.250	0.667	0.8	0.6	1.0	0.0	0.8	1.000	0.000	0.784	1.000	0.000	0.804
0.250	0.429	0.8	0.7	0.962	0.038	0.801	0.963	0.037	0.785	0.962	0.038	0.805
0.250	0.250	0.8	0.8	0.502	0.498	0.826	0.502	0.498	0.809	0.503	0.497	0.829
0.111	0.667	0.9	0.6	1.0	0.0	0.9	1.000	0.000	0.882	1.000	0.000	0.902
0.111	0.429	0.9	0.7	0.999	0.001	0.900	0.999	0.001	0.882	0.999	0.001	0.902
0.111	0.250	0.9	0.8	0.959	0.041	0.901	0.959	0.041	0.884	0.960	0.040	0.903
0.111	0.111	0.9	0.9	0.511	0.489	0.923	0.513	0.487	0.905	0.512	0.488	0.924

Table 7: Simulated data example, Bigamma distribution, without LoD

β_x^0	β_y^0	AUC_x	AUC_y	$\hat{\alpha}_{opt}$		$\widehat{AUC}$	α_{opt}^L		$\underline{AUC}$	α_{opt}^U		$\overline{AUC}$
$n_0 = n_1 = 50$, $\rho = 0$												
0.667	0.667	0.6	0.6	0.502	0.498	0.609	0.503	0.497	0.445	0.501	0.499	0.777
0.429	0.667	0.7	0.6	0.726	0.274	0.669	0.679	0.321	0.510	0.719	0.281	0.829
0.429	0.429	0.7	0.7	0.501	0.499	0.712	0.503	0.497	0.564	0.502	0.498	0.862
0.250	0.667	0.8	0.6	0.852	0.148	0.765	0.773	0.227	0.582	0.834	0.166	0.886
0.250	0.429	0.8	0.7	0.697	0.303	0.791	0.665	0.335	0.627	0.677	0.323	0.906
0.250	0.250	0.8	0.8	0.503	0.497	0.843	0.506	0.494	0.671	0.505	0.495	0.934
0.111	0.667	0.9	0.6	0.932	0.068	0.883	0.859	0.141	0.655	0.922	0.078	0.944
0.111	0.429	0.9	0.7	0.856	0.144	0.895	0.812	0.188	0.695	0.842	0.158	0.953
0.111	0.250	0.9	0.8	0.727	0.273	0.920	0.774	0.226	0.726	0.718	0.282	0.967
0.111	0.111	0.9	0.9	0.508	0.492	0.959	0.507	0.493	0.749	0.512	0.488	0.983
$n_0 = n_1 = 50$, $\rho = 0.5$												
0.667	0.667	0.6	0.6	0.523	0.477	0.599	0.522	0.478	0.471	0.526	0.474	0.746
0.429	0.667	0.7	0.6	0.826	0.174	0.659	0.779	0.221	0.535	0.818	0.182	0.802
0.429	0.429	0.7	0.7	0.515	0.485	0.680	0.520	0.480	0.576	0.514	0.486	0.824
0.250	0.667	0.8	0.6	0.949	0.051	0.762	0.868	0.132	0.608	0.927	0.073	0.869
0.250	0.429	0.8	0.7	0.808	0.192	0.761	0.728	0.272	0.641	0.774	0.226	0.874
0.250	0.250	0.8	0.8	0.516	0.484	0.795	0.519	0.481	0.682	0.514	0.486	0.899
0.111	0.667	0.9	0.6	0.987	0.013	0.886	0.919	0.081	0.682	0.975	0.025	0.938
0.111	0.429	0.9	0.7	0.952	0.048	0.879	0.857	0.143	0.715	0.929	0.071	0.937
0.111	0.250	0.9	0.8	0.831	0.169	0.884	0.790	0.210	0.743	0.803	0.197	0.943
0.111	0.111	0.9	0.9	0.520	0.480	0.921	0.525	0.475	0.772	0.523	0.477	0.962
$n_0 = n_1 = 50$, $\rho = 0.75$												
0.667	0.667	0.6	0.6	0.521	0.479	0.589	0.522	0.478	0.499	0.525	0.475	0.714
0.429	0.667	0.7	0.6	0.927	0.073	0.661	0.870	0.130	0.569	0.911	0.089	0.781
0.429	0.429	0.7	0.7	0.507	0.493	0.664	0.515	0.485	0.599	0.520	0.480	0.793
0.250	0.667	0.8	0.6	0.992	0.008	0.780	0.941	0.059	0.648	0.978	0.022	0.864
0.250	0.429	0.8	0.7	0.920	0.080	0.756	0.812	0.188	0.668	0.876	0.124	0.855
0.250	0.250	0.8	0.8	0.521	0.479	0.773	0.512	0.488	0.704	0.514	0.486	0.872
0.111	0.667	0.9	0.6	0.999	0.001	0.906	0.959	0.041	0.724	0.994	0.006	0.941
0.111	0.429	0.9	0.7	0.992	0.008	0.889	0.91	0.09	0.75	0.980	0.020	0.933
0.111	0.250	0.9	0.8	0.932	0.068	0.877	0.840	0.160	0.773	0.897	0.103	0.929
0.111	0.111	0.9	0.9	0.522	0.478	0.900	0.521	0.479	0.799	0.523	0.477	0.945
$n_0 = n_1 = 100$, $\rho = 0$												
0.667	0.667	0.6	0.6	0.502	0.498	0.609	0.503	0.497	0.445	0.501	0.499	0.777
0.429	0.667	0.7	0.6	0.726	0.274	0.669	0.679	0.321	0.510	0.719	0.281	0.829
0.429	0.429	0.7	0.7	0.501	0.499	0.712	0.503	0.497	0.564	0.502	0.498	0.862
0.250	0.667	0.8	0.6	0.852	0.148	0.765	0.773	0.227	0.582	0.834	0.166	0.886
0.250	0.429	0.8	0.7	0.697	0.303	0.791	0.665	0.335	0.627	0.677	0.323	0.906
0.250	0.250	0.8	0.8	0.503	0.497	0.843	0.506	0.494	0.671	0.505	0.495	0.934
0.111	0.667	0.9	0.6	0.932	0.068	0.883	0.859	0.141	0.655	0.922	0.078	0.944
0.111	0.429	0.9	0.7	0.856	0.144	0.895	0.812	0.188	0.695	0.842	0.158	0.953
0.111	0.250	0.9	0.8	0.727	0.273	0.920	0.774	0.226	0.726	0.718	0.282	0.967
0.111	0.111	0.9	0.9	0.508	0.492	0.959	0.507	0.493	0.749	0.512	0.488	0.983
$n_0 = n_1 = 100$, $\rho = 0.5$												
0.667	0.667	0.6	0.6	0.502	0.498	0.586	0.512	0.488	0.473	0.513	0.487	0.733
0.429	0.667	0.7	0.6	0.871	0.129	0.648	0.792	0.208	0.540	0.840	0.160	0.793
0.429	0.429	0.7	0.7	0.498	0.502	0.666	0.505	0.495	0.582	0.505	0.495	0.815
0.250	0.667	0.8	0.6	0.971	0.029	0.757	0.878	0.122	0.617	0.944	0.056	0.864
0.250	0.429	0.8	0.7	0.831	0.169	0.750	0.732	0.268	0.649	0.773	0.227	0.868
0.250	0.250	0.8	0.8	0.507	0.493	0.786	0.500	0.500	0.692	0.503	0.497	0.893
0.111	0.667	0.9	0.6	0.994	0.006	0.884	0.922	0.078	0.694	0.982	0.018	0.935
0.111	0.429	0.9	0.7	0.965	0.035	0.875	0.860	0.140	0.726	0.931	0.069	0.933
0.111	0.250	0.9	0.8	0.828	0.172	0.879	0.799	0.201	0.756	0.796	0.204	0.940
0.111	0.111	0.9	0.9	0.506	0.494	0.919	0.507	0.493	0.784	0.509	0.491	0.961
$n_0 = n_1 = 100$, $\rho = 0.75$												
0.667	0.667	0.6	0.6	0.503	0.497	0.576	0.509	0.491	0.502	0.513	0.487	0.701
0.429	0.667	0.7	0.6	0.970	0.030	0.655	0.897	0.103	0.577	0.936	0.064	0.774
0.429	0.429	0.7	0.7	0.505	0.495	0.651	0.504	0.496	0.607	0.503	0.497	0.784
0.250	0.667	0.8	0.6	0.999	0.001	0.778	0.954	0.046	0.660	0.989	0.011	0.860
0.250	0.429	0.8	0.7	0.955	0.045	0.750	0.816	0.184	0.679	0.889	0.111	0.848
0.250	0.250	0.8	0.8	0.507	0.493	0.765	0.502	0.498	0.715	0.498	0.502	0.867
0.111	0.667	0.9	0.6	1.000	0.000	0.906	0.962	0.038	0.738	0.997	0.003	0.939
0.111	0.429	0.9	0.7	0.998	0.002	0.888	0.912	0.088	0.763	0.986	0.014	0.930
0.111	0.250	0.9	0.8	0.949	0.051	0.874	0.844	0.156	0.787	0.900	0.100	0.926
0.111	0.111	0.9	0.9	0.512	0.488	0.899	0.506	0.494	0.812	0.514	0.486	0.943

Table 8: Simulated data example, Bigamma distribution, with LoD 10%

References

[1] Pepe MS, Thompson ML. Combining diagnostic test results to increase accuracy. *Biostatistics* 2000; **1**(2):123–140.

[2] McIntosh MW, Pepe MS. Combining several screening tests: Optimality of the risk score. *Biometrics* 2002; **58**(3):657–664.

[3] Ma S, Huang J. Combining multiple markers for classification using roc. *Biometrics* 2007; **63**(3):751–757.

[4] Pepe MS, Cai T, Longton G. Combining predictors for classification using the area under the receiver operating characteristic curve. *Biometrics* 2006; **62**(1):221–229.

[5] Kang L, Liu A, Tian L. Linear combination methods to improve diagnostic/prognostic accuracy on future observations. *Statistical Methods in Medical Research* 2013; .

[6] Su JQ, Liu JS. Linear combinations of multiple diagnostic markers. *Journal of the American Statistical Association* 1993; **88**(424):1350–1355.

[7] Liu A, Schisterman EF, Zhu Y. On linear combinations of biomarkers to improve diagnostic accuracy. *Statistics in Medicine* 2005; **24**:37–47.

[8] Perkins NJ, Schisterman EF, Vexler A. Receiver operating characteristic curve inference from a sample with a limit of detection. *American Journal of Epidemiology* 2007; **165**(3):325–333.

[9] Perkins NJ, Schisterman EF, Vexler A. Generalized roc curve inference for a biomarker subject to a limit of detection and measurement error. *Statistics in Medicine* 2009; **28**(13):1841–1860.

[10] Perkins NJ, Schisterman EF, Vexler A. Roc curve inference for best linear combination of two biomarkers subject to limits of detection. *Biometrical Journal* 2011; **53**(3):464–476.

[11] Dong T, Liu CC, Petricoin EF, Tang LL. Combining markers with and without the limit of detection. *Statistics in Medicine* 2014; **33**: 1307–1320.

[12] Augustin T, Coolen FPA. Nonparametric predictive inference and interval probability. *Journal of Statistical Planning and Inference* 2004; **124**(2):251–272.

[13] Coolen FPA. On nonparametric predictive inference and objective bayesianism. *Journal of Logic, Language and Information* 2006; **15**(1-2):21–47.

[14] Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford University Press: Oxford, 2003.

[15] Zhou XH, McClish DK, Obuchowski NA. *Statistical Methods in Diagnostic Medicine*. Wiley-Interscience: New York, 2002.

[16] Hill BM. Posterior distribution of percentiles: Bayes' theorem for sampling from a population. *Journal of the American Statistical Association* 1968; **63**(322):677–691.

[17] De Finetti B. *Theory of Probability: A Critical Introductory Treatment*. Wiley: London, 1974.

[18] Walley P. *Statistical Reasoning with Imprecise Probabilities*. Chapman & Hall: London, 1991.

[19] Weichselberger K. The theory of interval-probability as a unifying concept for uncertainty. *International Journal of Approximate Reasoning* 2000; **24**(2-3):149–170.

[20] Weichselberger K. *Elementare Grundbegriffe einer allgemeineren Wahrscheinlichkeitsrechnung I. Intervallwahrscheinlichkeit als umfassendes Konzept*. Physika: Heidelberg, 2001.

[21] Coolen-Maturi T, Coolen-Schrijner P, Coolen FPA. Nonparametric predictive inference for binary diagnostic tests. *Journal of Statistical Theory and Practice* 2012; **6**(5):665–680.

[22] Elkhafifi FF, Coolen FPA. Nonparametric predictive inference for accuracy of ordinal diagnostic tests. *Journal of Statistical Theory and Practice* 2012; **6**(4):681–697.

[23] Coolen-Maturi T, Coolen-Schrijner P, Coolen FPA. Nonparametric predictive inference for diagnostic accuracy. *Journal of Statistical Planning and Inference* 2012; **142**(5):1141–1150.

[24] Coolen-Maturi T, Elkhafifi FF, Coolen FP. Three-group roc analysis: A nonparametric predictive approach. *Computational Statistics & Data Analysis* 2014; **78**:69–81.

[25] Coolen-Maturi T. Three-group roc predictive analysis for ordinal outcomes. *Communications in Statistics Theory and Methods* 2016; To appear.

[26] Maturi TA, Coolen-Schrijner P, Coolen FPA. Nonparametric predictive pairwise comparison for real-valued data with terminated tails. *International Journal of Approximate Reasoning* 2009; **51**(1):141–150.

[27] Shafer G. *A Mathematical Theory of Evidence*. Princeton University Press: Princeton, NJ, 1976.

[28] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria 2016. URL `https://www.R-project.org/`.

[29] Bansal A, Pepe MS. When does combining markers improve classification performance and what are implications for practice? *Statistics in Medicine* 2013; **32**(11):1877–1892.

[30] Lawless JF, Fredette M. Frequentist prediction intervals and predictive distributions. *Biometrika* 2005; **92**:529–542.

[31] Coolen FPA. Nonparametric predictive inference. *International Encyclopedia of Statistical Science*, Lovric M (ed.). Springer, 2011; 968–970.

[32] Ma H, Bandos AI, Gur D. On the use of partial area under the ROC curve for comparison of two diagnostic tests. *Biometrical Journal* 2015; **57**(2):304–320.

[33] Wieand S, Gail MH, James BR, James KL. A family of nonparametric statistics for comparing diagnostic markers with paired or unpaired data. *Biometrika* 1989; **76**(3):585–592.

[34] Cox LH, Johnson MM, Kafadar K. Exposition of statistical graphics technology. *ASA Statistical Computing Section*, 1982; 55–56.

[35] Maturi TA. Nonparametric predictive inference for multiple comparisons. PhD Thesis, Durham University, Durham, UK 2010. Available from www.npi-statistics.com.

[36] Muhammad NB. Predictive inference with copulas for bivariate data. PhD Thesis, Durham University, Durham, UK 2016. Available from www.npi-statistics.com.

[37] Nakas CT, Yiannoutsos CT. Ordered multiple-class roc analysis with continuous measurements. *Statistics in Medicine* 2004; **23**(22):3437–3449.

[38] Nakas CT, Alonzo TA, Yiannoutsos CT. Accuracy and cut-off point selection in three-class classification problems using a generalization of the youden index. *Statistics in Medicine* 2010; **29**(28):2946–2955.

[39] Nakas CT, Dalrymple-Alford JC, Anderson T, Alonzo TA. Generalization of youden index for multiple-class classification problems applied to the assessment of externally validated cognition in parkinson disease screening. *Statistics in Medicine* 2013; **32**(6):995–1003.

[40] Ruopp MD, Perkins NJ, Whitcomb BW, Schisterman EF. Youden index and optimal cut-point estimated from observations affected by a lower limit of detection. *Biometrical Journal* 2008; **50**(3):419–430.